WORD HYPERBOLIC DEHN SURGERY

MARC LACKENBY

1. INTRODUCTION

In the late 1970's, Thurston dramatically changed the nature of 3-manifold theory with the introduction of his Geometrisation Conjecture, and by proving it in the case of Haken 3-manifolds [24]. The conjecture for general closed orientable 3-manifolds remains perhaps the most important unsolved problem in the subject. A weaker form of the conjecture [20] deals with the fundamental group of a closed orientable 3-manifold. It proposes that either it contains $\mathbb{Z} \oplus \mathbb{Z}$ as a subgroup or it is word hyperbolic, in the sense of Gromov [11]. Word hyperbolic groups are precisely those groups which satisfy a linear isoperimetric inequality. They are of fundamental importance in geometric group theory and have very many useful properties.

Many non-Haken 3-manifolds are known to satisfy the geometrisation conjecture, due to Thurston's hyperbolic Dehn surgery theorem [23]. This asserts that, if one starts with a compact orientable 3-manifold M with $M - \partial M$ supporting a complete finite volume hyperbolic structure, and one Dehn fills each component of ∂M, then one obtains a hyperbolic 3-manifold, providing that a finite number of 'exceptional' slopes are avoided on each component of ∂M. It remains an interesting unsolved problem to establish how large this collection of exceptional slopes can be. Hodgson and Kerckhoff have shown [13] that there is an upper bound on the number of exceptional slopes, the upper bound being a (large) number independent of M. In this paper, we will show that very many Dehn fillings yield a 3-manifold which is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group. We will approach this problem in a number of different ways.

Our first technique is differential geometric. We will establish an extension of Thurston and Gromov's 2π theorem [4], which we now describe. Pick a horoball neighbourhood N of the cusps of $M - \partial M$. Then, with respect to N, each slope s on ∂M inherits a *length* which is defined to be the length of the shortest curve

on ∂N with slope s. Thurston and Gromov showed that if the Dehn filling slope on each component of ∂M has length more than 2π, then the 3-manifold obtained by Dehn filling has a negatively curved Riemannian metric. This implies that it is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group. The following result improves the critical slope length from 2π to 6.

Theorem 3.1. *Let M be a compact orientable 3-manifold with interior having a complete finite volume hyperbolic structure. Let $s_1, \ldots, s_n$ be slopes on ∂M, with one s_i on each component of ∂M. Suppose that there is a horoball neighbourhood N of the cusps of $M - \partial M$ on which each s_i has length more than 6. Then, the manifold obtained by Dehn filling along $s_1, \ldots, s_n$ is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.*

This result has been established independently by Agol [2] (with the exception of word hyperbolicity, at this stage). An argument of Agol [2] gives the following corollary.

Corollary 3.4. *Let M be a compact orientable 3-manifold with a single torus boundary component, and with interior supporting a complete finite volume hyperbolic structure. Then all but at most 12 Dehn fillings on ∂M yield a 3-manifold which is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.*

Our second approach to word hyperbolic Dehn surgery is more combinatorial. We will define a structure which we call an 'angled ideal triangulation' on M, which was first studied by Casson. An *angled ideal triangulation* of M is an expression of $M - \partial M$ as a union of ideal 3-simplices glued along their faces, with each edge of each ideal 3-simplex having an associated interior dihedral angle in the range $(0, \pi)$. We insist that the three dihedral angles at each ideal vertex of each ideal 3-simplex sum to π, and that the sum of the dihedral angles around each edge is 2π.

When M has an angled ideal triangulation, each boundary component of M inherits a triangulation, with each triangle having angles summing to π. Define the length of an edge in ∂M to be $\min\{\alpha_1, \ldots, \alpha_6\}/2$, where $\alpha_1, \ldots, \alpha_6$ are the angles for the two triangles adjacent to that edge. Each simplicial path in ∂M

then inherits a length. We will define the 'combinatorial length' of a slope on ∂M, which will be *at least* the length of the shortest possible curve in the 1-skeleton of ∂M having that slope. Then we will prove (a more general version of) the following theorem, which is a combinatorial analogue of the 2π theorem.

Theorem 4.9. *Let M be a 3-manifold with an angled ideal triangulation. Let $s_1, \ldots, s_n$ be a collection of slopes on ∂M, with one s_i on each component of ∂M, and each with combinatorial length more than 2π. Then the manifold obtained by Dehn filling M along these slopes is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.*

Edge length =
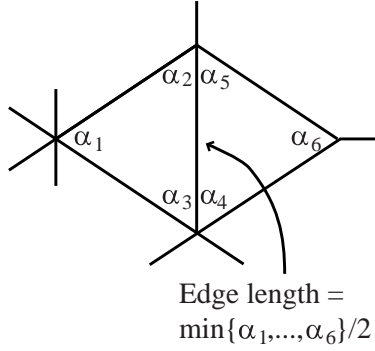$\min\{\alpha_1,...,\alpha_6\}/2$

Figure 1.

Any complete finite volume non-compact hyperbolic 3-manifold can be expressed as a union of convex hyperbolic ideal polyhedra glued isometrically along their faces [7]. This can usually be decomposed further into a hyperbolic ideal triangulation, which determines an angled ideal triangulation. However, it is *not* in general true that an angled ideal triangulation yields a complete hyperbolic structure. This is useful, since it implies that Mostow rigidity does not apply to angled ideal triangulations. One can therefore pick the most convenient angled ideal triangulation for one's purposes.

Our techniques extend beyond angled ideal triangulations to expressions of a 3-manifold as a union of ideal polyhedra with interior angles assigned to each edge. Dually, we will consider spines with certain extra combinatorial data. For example, the exterior of an alternating link admits a well-known spine arising from its diagram. Using this spine, we will deduce the following result about surgery along alternating links.

Let D be an alternating diagram of a knot or link L in S^3. We view D as a 4-valent graph $G(D)$ in S^2, equipped with 'under-over' crossing information. The diagram is *prime* if each simple closed curve in S^2 intersecting $G(D)$ transversely in two points in the interior of edges of $G(D)$ divides the 2-sphere into two discs, one of which contains no crossings of D. A region of D is a *bigon region* if it has precisely two edges of $G(D)$ in its boundary, and an edge of $G(D)$ is a *bigon edge* if it lies in the boundary of a bigon region. Define the *twist number $t(D)$* of D to be half the number of non-bigon edges of $G(D)$. For example, the alternating diagram D of the figure-eight knot in Figure 2 has twist number $t(D) = 2$. More generally, if K is a component of L, define $t(K, D)$ to be half the number of non-bigon edges lying in K. This number may be half-integral, but $t(D)$ is always an integer. Using this terminology, we can now state one of our main theorems.

Theorem 5.1. *Let D be a connected prime alternating diagram of a link L in S^3. For each component K of L, pick a surgery coefficient p/q (in its lowest terms) with $|q| > 8/t(K, D)$. The manifold obtained by Dehn surgery along L, via these surgery coefficients, is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.*

The figure eight knot

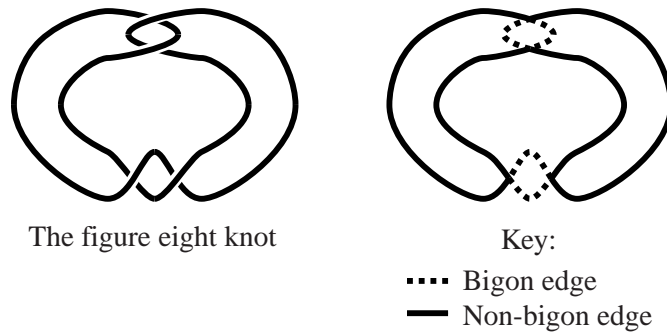Key:
···· Bigon edge
— Non-bigon edge

Figure 2.

Therefore, for 'sufficiently complicated' alternating links, every non-trivial surgery yields a 3-manifold satisfying the conclusions of the theorem. For example, this is true for all alternating knots with twist number at least 9. This has the following corollary.

Corollary 5.2. *There is a real number c with the property that if K is an alternating knot whose complement has a complete hyperbolic structure with volume at least c, then every non-trivial surgery on K yields a 3-manifold which is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.*

Surgery along alternating knots has been studied in a number of different contexts. Menasco and Thistlethwaite [19] embarked upon an analysis of surfaces with boundary in alternating knot exteriors, using combinatorial arguments in the spirit of [18]. They established that, under many circumstances, manifolds obtained by surgery along alternating knots do not contain embedded essential 2-spheres or tori. Their bounds on $|q|$ are a little stronger than those in Theorem 5.1, but their arguments only seem to work for embedded surfaces, and therefore do not imply that the fundamental group of the filled-in 3-manifold is infinite and word hyperbolic. Aitchison, Lumsden and Rubinstein [3] studied certain classes of alternating links and demonstrated the existence of non-positively curved metrics on their exteriors. Our approach is perhaps closest in spirit to theirs, but there is very little overlap in our results. Delman [6] and Roberts [22] have shown that every non-trivial surgery along a non-torus alternating knot yields a manifold with an essential lamination, and which is therefore irreducible and has infinite fundamental group [10]. Gabai and Kazez [9] have shown that closed atoroidal 3-manifolds with *genuine* essential laminations have word hyperbolic fundamental groups, but many of the laminations constructed by Roberts and Delman are not genuine.

Both our combinatorial approach and our differential geometric approach to word hyperbolic Dehn surgery have as their basis Gabai's Ubiquity theorem, which is a useful tool in showing that a 3-manifold has word hyperbolic fundamental group. In Section 2, we will offer a new simplified proof of this theorem. In Section 3, we will use this to prove Theorem 3.1, by examining 'geodesic spines' of non-compact complete hyperbolic 3-manifolds. In Section 4, we will examine angled spines of 3-manifolds (which are generalisations of angled ideal triangulations) and their relation to word hyperbolic Dehn surgery. In Section 5, we will apply these techniques to alternating links.

2. Gabai's Ubiquity theorem

Recall [11] that a finitely presented group G is *word hyperbolic* if, for some presentation $\langle g_1, \ldots, g_p \mid r_1, \ldots, r_q \rangle$ of G, a linear isoperimetric inequality is satisfied. In other words, there exists a constant $c \in \mathbb{R}$, such that, for every word w in $\langle g_1, \ldots, g_p \rangle$ representing the identity element of G, we can find an integer $n \leq c|w|$, where $|w|$ is the length of the word w, and a collection of elements $u_1, \ldots, u_n$ of $\langle g_1, \ldots, g_p \rangle$, such that

$$w = \prod_{i=1}^{n} u_i^{-1} r_{j_i} u_i.$$

Here, the above equality takes place in the free group $\langle g_1, \ldots, g_p \rangle$. It is not hard to see that if G satisfies a linear isoperimetric inequality for some presentation, then it will satisfy such an inequality for all presentations, although the choice of constant c might differ. If M is some compact manifold (possibly with boundary) with a Riemannian metric g, then its fundamental group is word hyperbolic if and only if, for some constant c, each homotopically trivial loop K in M bounds a disc D with

$$\text{Area}(D) \leq c \, \text{Length}(K).$$

Again, this property is independent of the Riemannian metric g, but the constant c is not. Similarly, we can consider a simplicial metric on M, and only consider curves K and discs D which are mapped into M simplicially. Alternatively, we can consider any metric on M which is bi-Lipschitz equivalent to a Riemannian metric.

Gabai introduced his Ubiquity theorem in [8]. He and Kazez used it in [9] to show that any closed atoroidal 3-manifold with a genuine essential lamination has word hyperbolic fundamental group. Here, we offer a simplified version of the theorem which is sufficient for our purposes, together with a new proof using cone manifolds. This proof has also been observed by Thurston. We say that a *slope* on a torus is the free homotopy class of an essential simple closed curve. If $s_1, \ldots, s_n$ are slopes on distinct toral boundary components of a 3-manifold M, then we denote the 3-manifold obtained by Dehn filling along these slopes by $M(s_1, \ldots, s_n)$.

Theorem 2.1. *Let M be a compact orientable 3-manifold whose interior supports a complete finite volume hyperbolic structure. Pick a Riemannian (respectively, simplicial) metric g on M. Let $s_1, \ldots, s_n$ be slopes on ∂M with one s_i on each component of ∂M, and suppose that the core of each surgery solid torus has infinite order in $\pi_1(M(s_1, \ldots, s_n))$. Suppose that for each smooth (respectively, simplicial) closed curve K in M which is homotopically trivial in $M(s_1, \ldots, s_n)$, there exists a map of a compact planar surface F into M with one component of ∂F being sent to K and each remaining component of ∂F being sent a non-zero multiple of one of the slopes $s_1, \ldots, s_n$, such that*

$$|F \cap \partial M| \leq c \, \mathrm{Length}(K, g).$$

Here, $c \in \mathbb{R}$ is a constant which may depend on M, g and $s_1, \ldots, s_n$, but which is independent of K and F. Then, $\pi_1(M(s_1, \ldots, s_n))$ is word hyperbolic.

The point behind the Ubiquity theorem is that, in order to establish the word hyperbolicity of $\pi_1(M(s_1, \ldots, s_n))$, one need not consider the area of discs in $M(s_1, \ldots, s_n)$, but can instead consider the number of times they intersect cores of the surgery solid tori.

Before we embark on the proof of this theorem, we need to establish some technical preliminaries. Let F be a compact surface mapped into a 3-manifold M. Then F is *homotopically ∂-incompressible* if no embedded essential arc R in F can be homotoped in M (keeping its endpoints fixed) to an arc in ∂M. Similarly, F is *homotopically incompressible* if the only simple closed curves in F which are homotopically trivial in M are those which bound discs in F.

Lemma 2.2. *Let M be a compact orientable 3-manifold with interior supporting a complete finite volume hyperbolic structure. Let F be a connected compact homotopically ∂-incompressible surface in M with negative Euler characteristic. Suppose that each component of ∂F either lies in ∂M, or is disjoint from ∂M and not homotopic to a curve in ∂M. Suppose at least one component of ∂F is of the former type, and at most one component of ∂F is of the latter type. Then there is a homotopy of $F - (\partial F \cap \partial M)$ to a pleated surface.*

Proof. This is well known, and more details can be found in [23]. Pick an ideal triangulation of $F - \partial F$, which is possible since F has negative Euler characteristic and non-empty boundary. We can ensure that no edge has both endpoints lying

in $\partial F - \partial M$. This determines a lamination on $F - (\partial F \cap \partial M)$ by 'spinning' the ideal triangulation of $F - \partial F$ around $\partial F - \partial M$ (if $\partial F - \partial M \neq \emptyset$) and then adding in $\partial F - \partial M$. There are (at most) three types of leaf in this lamination: the curve $\partial F - \partial M$; edges with both endpoints in $\partial F \cap \partial M$; and edges with an end spiralling towards $\partial F - \partial M$. The first of these can be homotoped to a geodesic, since it is not homotopic to a curve in ∂M. The second type of leaf can homotoped to a geodesic, since F is homotopically ∂-incompressible. Finally, the third type of leaf can be homotoped to a geodesic, since the endpoints of a lift of $\partial F - \partial M$ in $\mathbb{H}^3$ do not lie in the lift of a cusp of $M - \partial M$. The complement of the lamination is a collection of ideal triangles. Map each to a totally geodesic ideal triangle in $M - \partial M$. Then $F - \partial M$ is a pleated surface. □

Lemma 2.3. *Let F be a compact orientable incompressible ∂-incompressible surface properly embedded in a compact orientable 3-manifold M with incompressible boundary. Then F is homotopically incompressible and homotopically ∂-incompressible.*

Proof. This is a standard application of the Loop Theorem [12], which gives us immediately that F is homotopically incompressible. If F is homotopically ∂-compressible, then the doubled surface DF in the doubled manifold DM is not π_1-injective, and so is compressible, by the Loop Theorem. By considering outermost arcs of intersection between ∂M and a compressing disc for DF, we find that F is either compressible or ∂-compressible in M. □

Lemma 2.4. *Let F be a compact planar surface mapped into a compact 3-manifold M, with ∂M a union of tori. Let $s_1, \ldots, s_n$ be slopes on distinct components of ∂M. Suppose that all but one component of ∂F are sent to non-zero multiples of $s_1, \ldots, s_n$, but the remaining component C of ∂F lies on ∂M as a non-zero multiple of a slope other than $s_1, \ldots, s_n$. Then we may find such a surface which is homotopically incompressible and homotopically ∂-incompressible and which has no more boundary components than F.*

Proof. We will show that if F is homotopically compressible or homotopically ∂-compressible, then we can modify F to a new surface having the properties required of F but with fewer boundary components. This process will eventually terminate with the required surface. Suppose that there is an essential simple

closed curve R in F which is homotopically trivial in M. Then cut F along R and attach a disc to each copy of R, yielding a surface F'. The component of F' containing C has the required properties.

Suppose that there is an embedded essential arc R in F which can be homotoped (keeping ∂R fixed) to an arc R' in ∂M. Cut F along the arc R, and attach to the new surface two copies of the disc realising the homotopy between R and R', yielding a surface F'. There are a number of cases to consider.

Suppose first that the two points of ∂R lie in distinct components C_1 and C_2 of ∂F. Then these components are amalgamated into a single boundary component C' of F'. If neither C_1 nor C_2 is C, then C' is a multiple of one of the slopes $s_1, \ldots, s_n$. It may be homotopically trivial, in which case we cap it off with a disc. If one of C_1 and C_2 is C, then C' is a non-zero multiple of a slope other than $s_1, \ldots, s_n$. In each case, we have constructed the required surface.

Suppose now that the endpoints of R lie in a single component C_1 of ∂F. Then the arc R is separating in F (since F is planar). The boundary component C_1 becomes two boundary components C_1' and C_2' of $\partial F'$, which together are homologous in ∂M to C_1. If $C_1 = C$, then at least one of C_1' and C_2' (C_1', say) is a non-zero multiple of a slope other than $s_1, \ldots, s_n$. Taking the component of F' containing C_1' gives us the required surface. If $C_1 \neq C$, then there are two possibilities: either both C_1' and C_2' represent a multiple of one of the slopes $s_1, \ldots, s_n$, or neither do. In the former case, we take the component of F' containing C (possibly attaching a disc if one of its boundary components is homotopically trivial). In the latter case, we take the component of F' not containing C. □

Very similar arguments give the following result.

Lemma 2.5. *Let F be a compact planar surface mapped into a compact 3-manifold M, with ∂M a union of tori. Let $s_1, \ldots, s_n$ be slopes on distinct components of ∂M, and suppose that the core of each surgery solid torus in $M(s_1, \ldots, s_n)$ has infinite order in $\pi_1(M(s_1, \ldots, s_n))$. Suppose that all but one component of ∂F are sent to non-zero multiples of $s_1, \ldots, s_n$, but the remaining component of ∂F is sent to a knot K in the interior of M. Then we may find such a surface which is homotopically incompressible and homotopically ∂-incompressible and which has no more boundary components than F.*

Proof of Theorem 2.1. Let h be the complete finite volume hyperbolic metric on $M - \partial M$. If the inequality of the theorem holds for some Riemannian or simplicial metric g on M, it holds for every such metric, after possibly changing our choice of constant c. We are therefore free to give M the Riemannian metric obtained by removing from $(M - \partial M, h)$ a horoball neighbourhood N of its cusps. To check the word hyperbolicity of $\pi_1(M(s_1, \ldots, s_n))$, it suffices to verify a linear isoperimetric inequality for some metric on $M(s_1, \ldots, s_n)$ which is bi-Lipschitz equivalent to a Riemannian metric. We give $M(s_1, \ldots, s_n)$ a cone metric k which is a hyperbolic Riemannian metric away from the cores C of the surgery solid tori, but which may have cone angle other than 2π at C. That such a cone metric exists on $M(s_1, \ldots, s_n)$ is a well known consequence of the proof of Thurston's hyperbolic Dehn surgery theorem [23]. There is a map from $(M - \partial M, h)$ to $(M(s_1, \ldots, s_n), k)$ which does not increase distances by more than a factor of $c_1 > 1$ (say) and is, in fact, bi-Lipschitz with (constant c_1) away from N. It collapses a small horoball neighbourhood of the cusps (lying inside N) to the curves C.

We now consider a homotopically trivial curve K in $M(s_1, \ldots, s_n)$, and will construct a disc bounded by K satisfying a linear isoperimetric inequality. A homotopy (with area linearly bounded by $\text{Length}(K, k)$) pulls K away from a regular neighbourhood of C. We therefore view K as lying in $(M - \partial M) - N$. A further small homotopy takes K to a piecewise geodesic curve, each geodesic segment having length at least ϵ, say, where ϵ depends only on M.

If K is homotopically trivial in $M - \partial M$, K bounds a disc D in $M - \partial M$ with $\text{Area}(D, h) \leq c_2 \text{Length}(K, h)$, where $c_2 > 0$ depends only on M. This is because D may be realised as a union of at most $((\text{Length}(K, h)/\epsilon) - 2)$ geodesic triangles, each of which has area at most π. This implies that

$$\text{Area}(D, k) \leq c_1^2 \text{Area}(D, h) \leq c_1^2 c_2 \text{Length}(K, h) \leq c_1^3 c_2 \text{Length}(K, k),$$

which is the required linear isoperimetric inequality.

Suppose therefore that K is homotopically non-trivial in M. As in the statement of the theorem, there exists a map of a compact planar surface F into M, satisfying the inequality $|F \cap \partial M| \leq c \, \text{Length}(K, h)$. We may take such an F for which $|F \cap \partial M|$ is minimal, and so by Lemma 2.5, we may assume that F is homotopically ∂-incompressible. If $|F \cap \partial M| = 1$, then, as above, we let $F - (\partial F \cap \partial M)$

be a union of geodesic triangles and extend it to a disc D in $M(s_1, \ldots, s_n)$ with $\text{Area}(D, k) \leq c_1^3 c_2 \text{Length}(K, k)$. If K is homotopic in M to a curve in ∂M, then this curve is either a multiple of some s_i (in which case, we may take $|F \cap \partial M| = 1$) or is not a multiple of some s_i (in which case, a core of one of the surgery solid tori has finite order in $\pi_1(M(s_1, \ldots, s_n))$, contrary to assumption). Hence, we may assume that F satisfies the requirements of Lemma 2.2 and so we may homotope $F - (\partial F \cap \partial M)$ to a pleated surface F_1, say. This homotopy will take K to a geodesic K_1. It is well known [11] that the annulus A realising the free homotopy between K and K_1 can be taken to have

$$\text{Area}(A, h) \leq c_3 \text{ Length}(K, h),$$

for some constant $c_3 > 0$ which depends only on M (by an argument similar to the case where K is homotopically trivial in M). The pleated surface F_1 has

$$\text{Area}(F_1, h) = -2\pi\chi(F_1) = 2\pi(|F \cap \partial M| - 1) < 2\pi c \text{ Length}(K, h).$$

Therefore, by gluing F_1 and A, we get that K is part of the boundary of a surface F_2 as in the statement of the theorem with

$$\text{Area}(F_2, h) \leq (2\pi c + c_3)\text{Length}(K, h).$$

This surface F_2 extends to a disc D in $M(s_1, \ldots, s_n)$, with

$$\text{Area}(D, k) \leq c_1^2 (2\pi c + c_3)\text{Length}(K, h) \leq c_1^3 (2\pi c + c_3)\text{Length}(K, k),$$

which is the required isoperimetric inequality. □

3. An extension of the 2π theorem

The goal of this section is to prove the following result.

Theorem 3.1. *Let M be a compact orientable 3-manifold with interior having a complete finite volume hyperbolic structure. Let $s_1, \ldots, s_n$ be slopes on ∂M, with one s_i on each component of ∂M. Suppose that there is a horoball neighbourhood N of the cusps of $M - \partial M$ on which each s_i has length more than 6. Then, the manifold obtained by Dehn filling along $s_1, \ldots, s_n$ is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.*

Our first step is to construct a 'geodesic spine' of M. Recall that a *spine* of M is an embedded 2-dimensional cell complex S with the property that M is a regular neighbourhood of S. A *geodesic spine* is defined to be a spine of M in which each cell is totally geodesic. Define the *geodesic spine arising from the horoball neighbourhood N* to be

$$S = \{x \in (M - \partial M) - N : x \text{ does not have a unique closest point in } N\}.$$

To see that S is a geodesic spine, consider its inverse image $\tilde{S}$ under the covering map $\mathbb{H}^3 \to M - \partial M$. In this cover, N lifts to a collection of disjoint horoballs $\tilde{N}$. Let $\tilde{E}$ be the closure of a component of $\mathbb{H}^3 - \tilde{S}$. Each point in $\tilde{E} - \tilde{S}$ is closest to a single component of $\tilde{N}$. Placing this component of $\tilde{N}$ at ∞ in the upper half space model of $\mathbb{H}^3$, we see that there is a natural vertical retraction $\tilde{E} \to \partial \tilde{E}$. This retraction is defined for all components of $\mathbb{H}^3 - \tilde{S}$ and is equivariant with respect to the covering $\mathbb{H}^3 \to M - \partial M$. Hence, it descends to a retraction $M - \partial M \to S$. Consider a point of $\tilde{S}$. If $\{n_1, \ldots, n_t\}$ are its nearest points in $\tilde{N}$, then $n_1, \ldots, n_t$ lie in distinct components $\tilde{N}_1, \ldots, \tilde{N}_t$ of $\tilde{N}$. Also contained in $\tilde{S}$ are the set of points in $\mathbb{H}^3$ which are equidistant to $\tilde{N}_1, \ldots, \tilde{N}_t$ and not closer to any other component of $\tilde{N}$. This forms a totally geodesic convex subset of $\mathbb{H}^3$, which we take to be a single cell in our cell structure on $\tilde{S}$. This descends to a cell structure on S, which makes it into a geodesic spine.

Note that when M has a single boundary component, the geodesic spine is independent of the choice of horoball neighbourhood N of the cusp. However, when ∂M has more than one component, the geodesic spine need not be unique. Possibly the simplest example of this construction is the geodesic spine S for the

12

figure-eight knot complement, shown in Figure 3. A cross-section of $\tilde{S}$ in the upper half space model of $\mathbb{H}^3$ is shown in Figure 4.
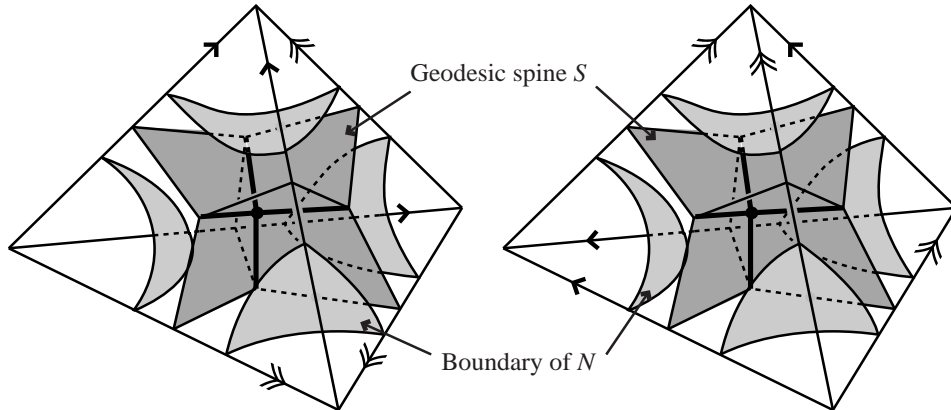

Geodesic spine S

Boundary of N

Figure 3.

$\tilde{N}$

$\tilde{S}$

$z=z_0$

$z=1$

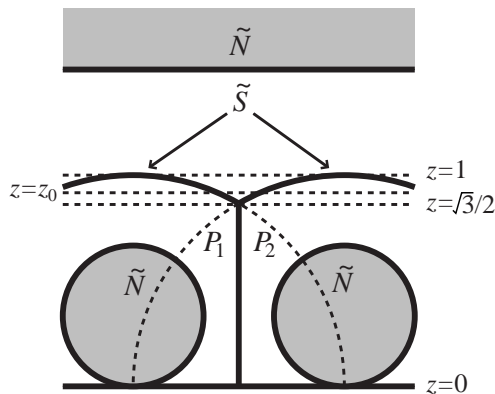$z=\sqrt{3}/2$

P_1 P_2

$\tilde{N}$ $\tilde{N}$

$z=0$

Figure 4.

An examination of S in this case will be instructive, since it satisfies a number of 'extremal' properties, which will be central to our arguments. Arrange $\tilde{N}$ so that it touches ∞ in the upper half space model for $\mathbb{H}^3$, and arrange $\tilde{S}$ so that it just touches $\{(x, y, z) \in \mathbb{H}^3 : z = 1\}$ (which we abbreviate to $\{z = 1\}$). Let z_0 be a real number in the range $(\sqrt{3}/2, 1)$. Note that $\tilde{S} \cap \{z = z_0\}$ is a collection of disjoint circles. Let $R(z_0)$ be the radius of each circle, and let $D(z_0)$ be the minimal distance between circles, measured in the Euclidean metric on $\{z = z_0\}$. Note that $2R(z_0) + D(z_0)$ is the distance between the centres of adjacent circles, which is $1/z_0$. The numbers $R(z_0)$ and $D(z_0)$ are extremal in the following sense.

Lemma 3.2. *Let M be a compact orientable 3-manifold whose interior supports a complete finite volume hyperbolic structure. Let N be a horoball neighbourhood of its cusps, and let S be the associated geodesic spine. Let $\tilde{N}$ and $\tilde{S}$ be the inverse images of N and S in $\mathbb{H}^3$. Arrange one component $\tilde{N}_0$ of $\tilde{N}$ as $\{z \geq 1\}$ in the upper half space model of $\mathbb{H}^3$, and let $\tilde{E}$ be the closure of the component of $\mathbb{H}^3 - \tilde{S}$ containing $\tilde{N}_0$. Let z_0 be a real number in the range $(\sqrt{3}/2, 1)$. Then the intersection $(\mathbb{H}^3 - \tilde{E}) \cap \{z = z_0\}$ is a (possibly empty) collection of disjoint discs. The radius of each disc is at most $R(z_0)$, and the distance between two discs is at least $D(z_0)$.*

Proof. Construct $\tilde{E}$ as follows. For each component $\tilde{N}_i$ of $\tilde{N}$ other than $\tilde{N}_0$, let P_i be the totally geodesic plane equidistant between $\tilde{N}_0$ and $\tilde{N}_i$. Then $\tilde{E}$ is the set of all points above $\bigcup_i P_i$ in the upper half space model. We will take two such planes P_1 and P_2 and will move them in $\mathbb{H}^3$ using the sequence of operations in Figure 5 until they are the planes P_1 and P_2 of Figure 4. Each operation will not decrease the radius of $P_i \cap \{z = z_0\}$, nor will it increase the distance between $P_1 \cap \{z = z_0\}$ and $P_2 \cap \{z = z_0\}$. Since these circles end up as being disjoint, they must have started disjoint. Thus, $(\mathbb{H}^3 - \tilde{E}) \cap \{z = z_0\}$ must have started as a (possibly empty) collection of disjoint discs, satisfying the required properties.

The first operation is to translate $\tilde{N}_1$ and $\tilde{N}_2$ horizontally until they touch. The second operation is to scale (in the Euclidean metric on the upper half space model) both $\tilde{N}_1$ and $\tilde{N}_2$ by the same factor, keeping them just touching one another, until at least one of them ($\tilde{N}_2$, say) also touches $\tilde{N}_0$. We can perform this scale so that the set of points above the old $\tilde{N}_1 \cup \tilde{N}_2$ contains the set of points above the new $\tilde{N}_1 \cup \tilde{N}_2$. This implies that the old $P_1 \cup P_2$ lies below the new $P_1 \cup P_2$, since P_i is the set of points equidistant between $\tilde{N}_0$ and $\tilde{N}_i$. Hence, we have not decreased the radii of $P_i \cap \{z = z_0\}$, nor have we moved $P_1 \cap \{z = z_0\}$ and $P_2 \cap \{z = z_0\}$ further from each other. This is also true of the third operation, which leaves $\tilde{N}_2$ unchanged but expands and translates $\tilde{N}_1$, keeping $\tilde{N}_2$ and $\tilde{N}_1$ just touching each other. We end with $\tilde{N}_1$ and $\tilde{N}_2$ as in Figure 5 and hence P_1 and P_2 as in Figure 4. $\square$
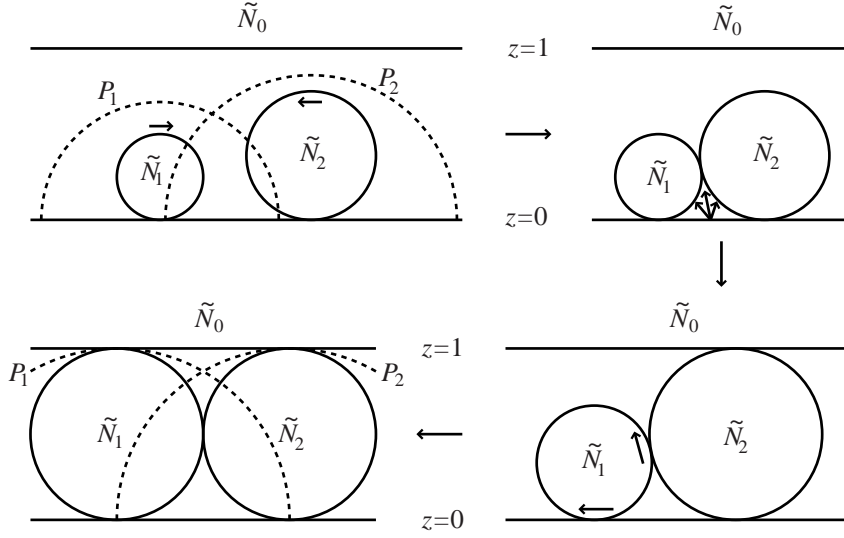
Figure 5.

The basis behind Theorem 3.1 is the following area estimate. The factor $(\pi/3)$ in this estimate is the reason for the improvement of the critical slope length from 2π to 6.

Lemma 3.3. *Let M be a compact orientable 3-manifold whose interior supports a complete finite volume hyperbolic structure. Let S be a geodesic spine arising from a horoball neighbourhood N of the cusps of $M - \partial M$. Let G be a compact orientable (possibly non-embedded) surface in the closure of a component of $M - S$, with $\partial G \subset \partial M \cup S$ and with $\partial G \cap \partial M$ a single curve representing $\pm k[s] \in H_1(\partial M)$, where $k \in \mathbb{N}$ and s is some slope. Then*

$$\mathrm{Area}(G - \partial M) \geq k(\pi/3)\,\mathrm{Length}(s),$$

where the slope length of s is measured with respect to N.

Proof. Let E be the closure of the component of $(M - \partial M) - S$ which contains $G - \partial M$, and let $\tilde{E}$ be a component of the inverse images of E in $\mathbb{H}^3$. As in Lemma 3.2, arrange the component $\tilde{N}_0$ of $\tilde{N}$ lying in $\tilde{E}$ as $\{z \geq 1\}$. For $z_0 \geq \sqrt{3}/2$, let $E(z_0)$ be the image of $\tilde{E}(z_0) = \tilde{E} \cap \{z \geq z_0\}$ under the covering map $\mathbb{H}^3 \to M - \partial M$. We may assume (after a very small homotopy of G) that, for all but finitely many z_0, $\partial E(z_0) \cap G$ is a collection C of immersed curves in $\partial E(z_0)$ which is homologous to $\pm k[s]$ in ∂M. For each component of C, pick a lift to an arc or closed curve in $\partial \tilde{E}(z_0)$, each arc having endpoints on $\tilde{S}$, and let $\tilde{C}'$ be the union of these lifts over

15

all components of C. Let $\tilde{C}$ be the image of $\tilde{C}'$ under the vertical projection map $\partial\tilde{E}(z_0) \to \{z = z_0\}$. Discard any components of $\tilde{C}$ disjoint from $\tilde{E}$. Straighten $\tilde{C}$ inside each disc component of $(\mathbb{H}^3 - \tilde{E}) \cap \{z = z_0\}$, creating arcs and circles $\tilde{C}(z_0)$, say, with total length at least $k\,\mathrm{Length}(s)/z_0$. (See Figure 6.)
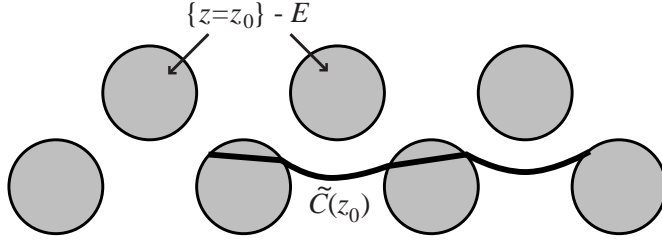

Figure 6.

For our purposes, what is relevant is the ratio of $\mathrm{Length}(\tilde{C}(z_0) \cap \tilde{E})$ to $\mathrm{Length}(\tilde{C}(z_0) - \tilde{E})$. Suppose initially that no component of $\tilde{C}(z_0)$ lies entirely in $\tilde{E}$. The arcs of $\tilde{C}(z_0) \cap \tilde{E}$ alternate with the arcs of $\tilde{C}(z_0) - \tilde{E}$. By Lemma 3.2, each of the former arcs has length at least $D(z_0)$, and each of the latter arcs has length at most $2R(z_0)$. Hence,

$$\mathrm{Length}(\tilde{C}(z_0) \cap \tilde{E}) \geq \frac{D(z_0)}{D(z_0) + 2R(z_0)}\mathrm{Length}(\tilde{C}(z_0)) \geq D(z_0)k\,\mathrm{Length}(s).$$

Note that this is true even if components of $\tilde{C}(z_0)$ lie entirely in $\tilde{E}$. It also remains true for $z_0 \geq 1$ if we define $D(z_0) = 1/z_0$ and $R(z_0) = 0$. Hence

$$\mathrm{Area}(G - \partial M) \geq \int_{\sqrt{3}/2}^{\infty} \frac{\mathrm{Length}(\tilde{C}(z) \cap \tilde{E})}{z}\,dz$$
$$\geq k\,\mathrm{Length}(s)\int_{\sqrt{3}/2}^{\infty} \frac{D(z)}{z}\,dz.$$

But, it is clear from Figure 7, that the final integral is precisely one third of the area of an ideal triangle. Therefore $\mathrm{Area}(G - \partial M) \geq k(\pi/3)\,\mathrm{Length}(s)$. $\square$
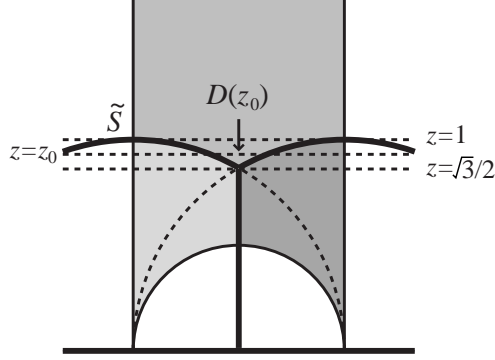
16

Figure 7.

Proof of Theorem 3.1. Note first that if $M(s_1, \ldots, s_n)$ is irreducible and has infinite, word hyperbolic fundamental group, then it is neither toroidal nor Seifert fibred. Suppose that $M(s_1, \ldots, s_n)$ is reducible. Then we can find a compact, incompressible, ∂-incompressible, planar surface F properly embedded in M, with each component of ∂F having one of the slopes $s_1, \ldots, s_n$, and with $|\partial F| \geq 3$. By Lemmas 2.3 and 2.2, we may homotope $F - \partial F$ to a pleated surface in $M - \partial M$. It then inherits a hyperbolic Riemannian metric which has

$$\text{Area}(F - \partial F) = -2\pi\chi(F) = 2\pi(|F \cap \partial M| - 2) < 2\pi|F \cap \partial M|.$$

Each component of intersection between F and $M - S$ intersects ∂M at most once, otherwise F would be ∂-compressible in M. Hence, Lemma 3.3 gives that

$$\text{Area}(F - \partial F) \geq |F \cap \partial M|(\pi/3) \min_{1 \leq i \leq n} \text{Length}(s_i) > 2\pi|F \cap \partial M|,$$

which gives a contradiction.

Similarly, we claim that the core of each surgery solid torus in $M(s_1, \ldots, s_n)$ has infinite order in $\pi_1(M(s_1, \ldots, s_n))$. In particular, $\pi_1(M(s_1, \ldots, s_n))$ is infinite. For, if not, some non-zero multiple of some core curve forms the boundary of a disc mapped into $M(s_1, \ldots, s_n)$. We can ensure that the intersection of this disc with M is a compact planar surface F mapped into M with all but one component of ∂F having slope a non-zero multiple of some s_i, and the remaining component sent to a non-zero multiple of a slope other than $s_1, \ldots, s_n$. By Lemma 2.4, we may take F to be homotopically ∂-incompressible, and so we may homotope $F - \partial F$ into pleated form. The above argument gives us a contradiction.

In order to prove that $\pi_1(M(s_1, \ldots, s_n))$ is word hyperbolic, we will use Gabai's Ubiquity theorem. We need to give M some metric; this will be the

17

hyperbolic Riemannian metric on $(M - \partial M) - \text{int}(N)$. Consider a curve K in $(M - \partial M) - \text{int}(N)$ which is homotopically trivial in $M(s_1, \ldots, s_n)$. There exists a compact planar surface F in M, with ∂F consisting of K and curves on ∂M, each representing a non-zero multiple of one of the slopes $s_1, \ldots, s_n$. Since each surgery solid torus in $M(s_1, \ldots, s_n)$ has infinite order in $\pi_1(M(s_1, \ldots, s_n))$, Lemma 2.5 implies that we may take F to be homotopically ∂-incompressible. We will show that there is a constant c and a choice of F with $|F \cap \partial M| \leq c \, \text{Length}(K)$. If K is homotopically trivial in M, then this is immediate. If K is homotopically non-trivial, then there is a lower bound on its length. Hence, by taking c sufficiently large, we may assume that $|F \cap \partial M| \geq 2$. Hence, by Lemma 2.2, we may homotope $F - \partial M$ in $M - \partial M$ to a pleated surface F_1, taking K to geodesic K_1. Since F_1 is pleated, its area is

$$\text{Area}(F_1) = -2\pi\chi(F_1) < 2\pi|F \cap \partial M|.$$

As in the proof of Theorem 2.1, we may ensure that the annulus realising the free homotopy between K and K_1 has area at most $c_3\text{Length}(K)$, where $c_3 > 0$ is a constant depending only on M. Gluing this annulus to F_1 gives a surface F_2 with

$$\text{Area}(F_2) < 2\pi|F \cap \partial M| + c_3\text{Length}(K).$$

Let E be the closure of some component of $(M - \partial M) - S$. For any $z_0 \geq \sqrt{3}/2$, define $E(z_0)$ as in the proof of Lemma 3.2. Then, there is a vertical projection $E(\sqrt{3}/2) \to \partial E(\sqrt{3}/2)$. For each component A of $K \cap E(\sqrt{3}/2)$, let $D(A)$ be the vertical disc in $E(\sqrt{3}/2)$ lying below it. Since K is disjoint from $\text{int}(N)$, A lies in the region $\{\sqrt{3}/2 \leq z \leq 1\}$, and so

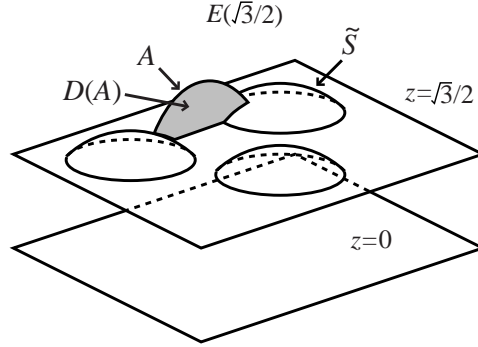$$\text{Area}(D(A)) \leq (2/\sqrt{3} - 1)\text{Length}(A).$$

$E(\sqrt{3}/2)$

A

$\widetilde{S}$

$D(A)$

$z=\sqrt{3}/2$

$z=0$

Figure 8.

Let G be a component of $F_2 - S$ which touches ∂M. Since F is homotopically ∂-incompressible, G must have a single boundary component which represents $\pm k[s_j] \in H_1(\partial M)$, for some $s_j \in \{s_1, \ldots, s_n\}$ and $k \in \mathbb{N}$. If G is not disjoint from K, let $A_1, \ldots, A_t$ be the arcs of $K \cap G$. Attach $D(A_1) \cup \ldots \cup D(A_t)$ to G to form a surface G_1. (If G is disjoint from K, let $G_1 = G$.) By Lemma 3.3,

$$\text{Area}(G_1) \geq k(\pi/3) \, \text{Length}(s_j) \geq (\pi/3) \, \text{Length}(s_j),$$

and so

$$\text{Area}(G) \geq (\pi/3) \, \text{Length}(s_j) - (2/\sqrt{3} - 1)\text{Length}(K \cap G).$$

Summing these inequalities over each component G of $F_2 - S$ which touches ∂M, we get that

$$\text{Area}(F_2) \geq |F \cap \partial M|(\pi/3) \min_{1 \leq j \leq n} \text{Length}(s_j) - (2/\sqrt{3} - 1)\text{Length}(K).$$

Therefore,

$$2\pi|F \cap \partial M| + c_3\text{Length}(K) > \text{Area}(F_2)$$
$$\geq |F \cap \partial M|(\pi/3) \min_{1 \leq j \leq n} \text{Length}(s_j) - (2/\sqrt{3} - 1)\text{Length}(K),$$

which gives that

$$|F \cap \partial M|\left[\left((\pi/3) \min_{1 \leq j \leq n} \text{Length}(s_j)\right) - 2\pi\right] \leq (2/\sqrt{3} - 1 + c_3)\text{Length}(K).$$

Theorem 2.1 now gives that $\pi_1(M(s_1, \ldots, s_n))$ is word hyperbolic. $\square$

Corollary 3.4. *Let M be a compact orientable 3-manifold with a single torus boundary component, and with interior supporting a complete finite volume hyperbolic structure. Then all but at most 12 Dehn fillings on ∂M yield a 3-manifold*

19

which is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.

Proof. It suffices to show that at most 12 slopes on ∂M have length no more than 6 (with respect to the maximal horoball neighbourhood N of ∂M). This was established by Agol [2], using a new result of Cao and Meyerhoff [5]. Here is a brief outline of Agol's argument. He showed first that if two slopes s_1 and s_2 on ∂M have length at most six, then their distance $\Delta(s_1, s_2)$ on ∂M is at most 10. This is because a simple and well-known argument in Euclidean geometry gives that

$$\Delta(s_1, s_2) \leq \frac{\text{Length}(s_1) \, \text{Length}(s_2)}{\text{Area}(\partial N)},$$

and Cao and Meyerhoff have recently shown that $\text{Area}(\partial N) \geq 3.35$. This gives that $\Delta(s_1, s_2) \leq 10$. Agol then showed that if $\mathcal{S}$ is a collection of distinct slopes on a torus, with any two elements of $\mathcal{S}$ having distance at most 10, then $|\mathcal{S}| \leq 12$. □

It is worth noting that, in many cases, the bounds of Theorem 3.1 will be far from optimal. The basis behind Theorem 3.1 is the area estimate in Lemma 3.3, which exploits the fact that the area of the surface G embedded in the complement of the geodesic spine S picks up more area than the parts of G lying in the horoball neighbourhood N. If the manifold $M - \partial M$ has large volume in comparison with that of N, then there will be parts of S lying far from N. If G approaches these parts of S, then its area will be greater than Lemma 3.3 predicts.

4. Angled spines of 3-manifolds

Casson has realised that useful results may deduced about non-compact finite volume complete hyperbolic 3-manifolds merely by examining the interior dihedral angles of a straight ideal triangulation. He studied closed normal surfaces in these ideal triangulations, and he showed that normal 2-spheres do not occur, and that the only normal tori are links of the ideal vertices. This section is devoted to showing that Casson's ideas may be extended to surfaces with boundary, and in this way, we will deduce some interesting results about Dehn surgery.

Recall that an *ideal triangulation* of a compact 3-manifold M with non-empty boundary is an expression of $M - \partial M$ as a union of 3-simplices glued along their faces, with the vertices then removed. An *angled ideal triangulation* is an ideal triangulation, together with the assignment of a real number in the range $(0, \pi)$ to each edge of each 3-simplex, known as the *interior angle* of that edge. We insist that the interior angles around each edge sum to 2π, and that the sum of the three interior angles at each vertex of each 3-simplex is π.

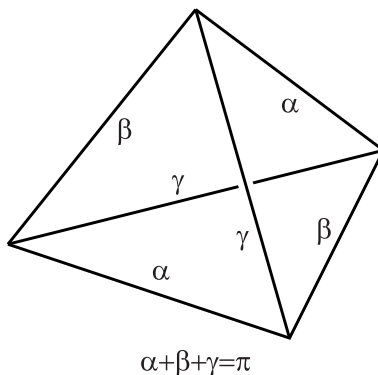

$\alpha+\beta+\gamma=\pi$

Figure 9.

The dual picture of an ideal triangulation is a special spine. If the ideal triangulation is angled, then the associated special spine inherits a certain combinatorial structure. We generalise this below, by considering spines more general than special spines. A spine of a 3-manifold M with non-empty boundary can be thickened to a handle decomposition $\mathcal{H}$ of M. Hereafter, we will only consider handle structures arising in this way. We will denote the i-handles of $\mathcal{H}$ by $\mathcal{H}^i$. Particularly important will be the surface $\mathcal{H}^0 \cap (\mathcal{H}^1 \cup \mathcal{H}^2)$, which we will denote by $\mathcal{F}$. Providing each 2-handle of $\mathcal{H}$ touches some 1-handle, this surface $\mathcal{F}$ inherits a

handle structure, with the 0-handles $\mathcal{F}^0$ of $\mathcal{F}$ being $\mathcal{H}^0 \cap \mathcal{H}^1$, and the 1-handles $\mathcal{F}^1$ of $\mathcal{F}$ being $\mathcal{H}^0 \cap \mathcal{H}^2$. We will insist that $H_0 \cap \mathcal{F}$ is connected for each 0-handle H_0 of $\mathcal{H}$. When $\mathcal{H}$ is dual to an angled ideal triangulation, then each component of $\mathcal{F}$ is of the form shown in Figure 10 with 'interior' angles assigned to each 1-handle of $\mathcal{F}$. It is this structure which we wish to generalise.
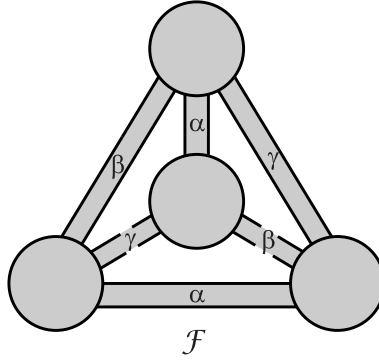

Figure 10.

There is a well-established theory [15] of normal embedded surfaces in a handle structure $\mathcal{H}$ of a 3-manifold M, which we now outline. Any incompressible surface F properly embedded in an irreducible 3-manifold M can be ambient isotoped so that afterwards, it intersects each handle of $\mathcal{H}$ in a collection of disjoint discs which respect the product structure on that handle, as in Figure 11. This gives the surface a handle structure, with the i-handles of M giving rise to the i-handles of F.
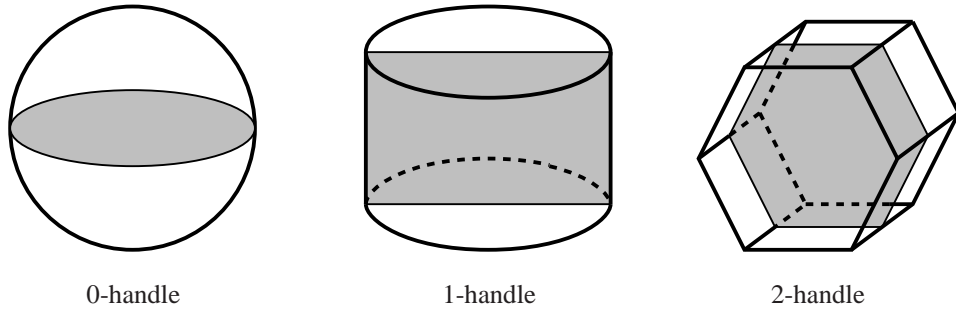

0-handle   1-handle   2-handle

Figure 11.

Furthermore, if F is ∂-incompressible and no component of F is a disc parallel to a disc in ∂M, then we can ensure that each curve N of $F \cap \partial \mathcal{H}^0$ is *normal*, which means that it is a simple closed curve satisfying each of the following conditions:

 (i) N respects the product structures of the 1-handles of $\mathcal{F}$;

 (ii) N does not lie entirely in ∂M or entirely in $\mathcal{F}^0$;

 (iii) no arc of intersection between N and $\mathcal{F}^0$ has endpoints lying in the same component of $\partial \mathcal{F}^0 \cap \mathcal{F}^1$, or in the same component of $\partial \mathcal{F}^0 - \mathcal{F}^1$, or in adjacent components of $\partial \mathcal{F}^0 \cap \mathcal{F}^1$ and $\partial \mathcal{F}^0 - \mathcal{F}^1$;

 (iv) no arc of intersection between N and $\partial \mathcal{H}^0 - \mathcal{F}$ has endpoints lying in the same component of $\partial \mathcal{F}^0 - \mathcal{F}^1$;

 (v) N intersects any given 1-handle of $\mathcal{F}$ in at most one arc;

 (vi) N intersects any given component of $\partial \mathcal{H}^0 - \mathcal{F}$ in at most one arc.

We will say that a surface F satisfying all of the above restrictions is *normal*.

These cannot be part of the
same normal curve N, by (v)

$\mathcal{F}$

Forbidden arcs
of N, by (iii)

These cannot be part of the
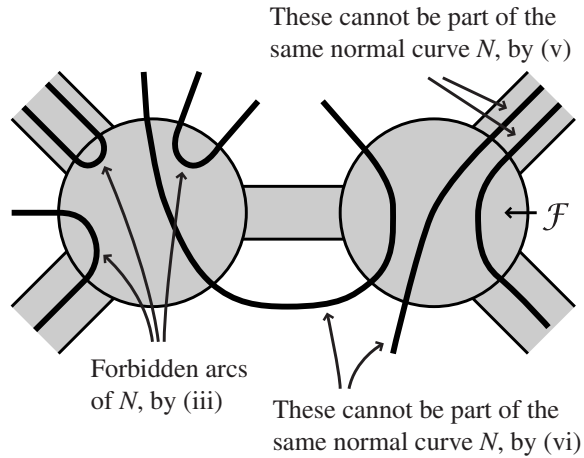same normal curve N, by (vi)

Figure 12.

We can now frame some crucial definitions. Consider a spine of a 3-manifold M with non-empty boundary. Assign two real numbers to each 1-handle of $\mathcal{F}$, each number lying in the range $(0, \pi)$, and the two numbers summing to π. These numbers are known as the *interior* and *exterior angles* of that 1-handle. If N is a normal curve, suppose that $\epsilon_1, \ldots, \epsilon_n$ are the exterior angles of the 1-handles of

$\mathcal{F}$ along which N runs. Then define the *combinatorial area* of N to be

$$a(N) = (\sum_{i=1}^{n} \epsilon_i) - 2\pi + \pi|N \cap \partial M|.$$

We will say that the angles on $\mathcal{F}^1$ determine an *angled spine* providing that

- every normal curve in $\partial \mathcal{H}^0$ has non-negative combinatorial area, and

- the interior angles around any 2-handle of $\mathcal{H}$ sum to 2π.

We define the *combinatorial area* $a(H)$ of a 0-handle H of a normal surface F to be the combinatorial area of ∂H. We also define the *combinatorial area* $a(F)$ of F to be the sum of the combinatorial areas of its 0-handles.

The reason for the above terminology is an analogy from hyperbolic geometry. In that context, the area of a planar convex hyperbolic polygon is the sum of its exterior angles, with 2π then subtracted. In the above formula, there is an extra term involving the number of intersections between N and ∂M. Again, this can be understood intuitively from hyperbolic geometry. If $M - \partial M$ is hyperbolic, then the parts of F approaching a cusp of $M - \partial M$ most naturally inherit zero interior angle there.

Note that there is only a finite number of normal curves in any given 0-handle of $\mathcal{H}$ up to ambient isotopy leaving $\mathcal{H}$ invariant. Thus, it is a simple process to check whether a given assignment of angles actually gives rise to an angled spine. Note also that the condition that a particular choice of exterior angles gives rise to an angled spine is a 'convex condition'. In other words, given two choices of angles, each of which yields an angled spine, then any convex linear combination of these angles also gives an angled spine. The following is a particularly important example of an angled spine.
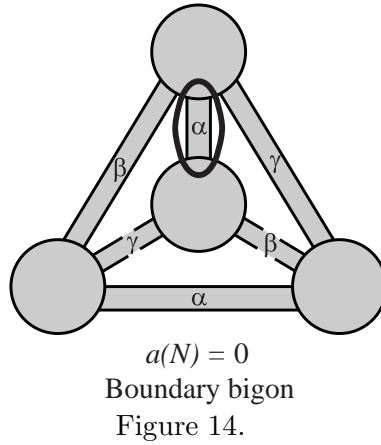
Lemma 4.1. *An angled ideal triangulation determines an angled spine.*

Proof. The interior angles around each edge of the ideal triangulation sum to 2π, which gives the second condition in the definition of an angled spine. To see the first condition, observe that each component of $\mathcal{F}$ is as shown in Figure 10. Each normal curve N which intersects ∂M in more than one arc automatically has non-negative combinatorial area. The normal curves intersecting ∂M in at most one arc are given in Figure 13 (up to the obvious symmetries of $\mathcal{F}$), and can

be seen to have non-negative combinatorial area. □

$a(N) = 0$
Triangle

$a(N) = 2\alpha$

$a(N) = \alpha$
Figure 13.
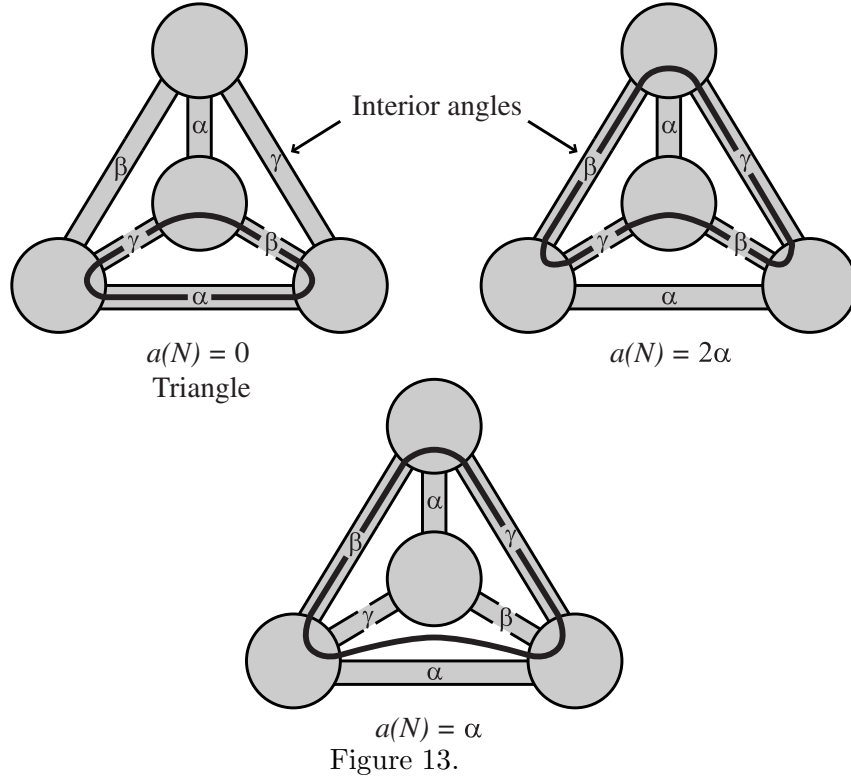
The above proof has the following corollary. Let $\mathcal{H}$ be a handle structure dual to an angled ideal triangulation. Then the only normal curves with zero combinatorial area in a 0-handle of $\mathcal{H}$ are *triangles* (as in the leftmost diagram of Figure 13) and *boundary bigons*, one of which is shown in Figure 14.

$a(N) = 0$
Boundary bigon
Figure 14.

More generally, we will say that a normal curve N in the boundary of a 0-handle

25

of an angled spine (not necessarily arising from an angled ideal triangulation) is a *boundary bigon* if it is disjoint from $\mathcal{F}^1$ and encloses a disc in $\partial \mathcal{H}^0$ containing a single 1-handle of $\mathcal{F}$.

We wish to consider surfaces F in M which need not be embedded; in this case, the theory of normal surfaces is less well understood. In fact, we will consider surfaces F which may even have boundary components not lying in ∂M. But, we will insist that ∂F is disjoint from $\mathcal{H}^2$, that its intersection with $\mathcal{H}^1$ respects the product structure on $\mathcal{H}^1$, and that if a component of ∂F touches ∂M then it lies entirely in ∂M. It is easy to see that any such surface F can be ambient isotoped (keeping ∂F invariant) so that, afterwards, its intersection with any i-handle $D^i \times D^{3-i}$ is of the form $D^i \times G$, for some codimension one submanifold G of D^{3-i}. After a further homotopy (keeping ∂F fixed) we may assume that F intersects each handle in a collection of discs, providing that F is homotopically incompressible and has no 2-sphere components, and M is irreducible. Each 0-handle H of F intersects $\partial \mathcal{H}^0$ in a 1-manifold. This 1-manifold is a collection of arcs if H touches $\partial F - \partial M$, but otherwise it is a closed curve. It need not be embedded, but we can ensure (after a further homotopy keeping $\partial F - \partial M$ fixed) that each arc of intersection with $\mathcal{F}^0$, $\mathcal{F}^1$ and $\partial \mathcal{H}^0 - \mathcal{F}$ is embedded. If F is homotopically ∂-incompressible and no component is a disc parallel to a disc in ∂M, and ∂M is incompressible, then we may further homotope F (keeping $\partial F - \partial M$ fixed) so that each curve of $F \cap \partial \mathcal{H}^0$ satisfies conditions (i) - (iv) in the above definition of normality (with the possible exception of arc components of $F \cap \partial \mathcal{H}^0$ which are allowed to lie entirely in $\mathcal{F}^0$). We say that F is *admissible* if it satisfies each of the above restrictions. Also, we say that a curve N in $\partial \mathcal{H}^0$ is *admissible* if it satisfies conditions (i) - (iv) and each arc of intersection with $\mathcal{F}^0$, $\mathcal{F}^1$ and $\partial \mathcal{H}^0 - \mathcal{F}$ is embedded. We will not insist that (v) and (vi) hold.

We wish to define the combinatorial area of an admissible surface F. Let H be a 0-handle of F, and, as above, let $\epsilon_1, \ldots, \epsilon_n$ be the exterior angles of the 1-handles of $\mathcal{F}$ over which ∂H runs (counted with multiplicity if it runs over a 1-handle of $\mathcal{F}$ more than once). Then define the *combinatorial area* of H to be

$$a(H) = (\sum_{i=1}^{n} \epsilon_i) - 2\pi + \pi|H \cap \partial M| + 3\pi|\partial H - \partial \mathcal{H}^0|.$$

The coefficient 3π is somewhat arbitrary; it was chosen so that the combinatorial

area of H is automatically positive if $\partial H \cap \partial \mathcal{H}^0$ is not a closed curve. Define the *combinatorial area* of F to be the sum of the combinatorial areas of its 0-handles. For combinatorial area to have useful applications to non-embedded surfaces, we need the following fact.

Lemma 4.2. *Let $\mathcal{H}$ be a thickened angled spine. Then each closed admissible curve N in $\partial \mathcal{H}^0$ has non-negative combinatorial area. Furthermore, if N is not a normal curve, then its combinatorial area is positive.*

Proof. We will prove this by induction on the number of arcs of intersection between N and $\mathcal{F}^1$. If N is normal then, from the definition of an angled spine, its combinatorial area is non-negative. Suppose therefore that N is not normal and that its combinatorial area is non-positive.

The induction starts with N disjoint from $\mathcal{F}^1$. Since N is not normal, it must have at least two arcs of intersection with ∂M. This gives it non-negative area. Since we are assuming that the area of N is non-positive, it must be zero. Hence, N has precisely two arcs of intersection with ∂M. If N is non-embedded in some 0-handle of $\mathcal{F}$, then it is not hard to construct a normal curve which intersects ∂M in a single arc and which is disjoint from $\mathcal{F}^1$. But this normal curve has negative combinatorial area, contradicting the assumption that this is an angled spine. Hence, the only way that N can fail to be normal if it is non-embedded in ∂M or violates condition (vi) of the definition of normality. Again, we can construct a negative area normal curve in $\partial \mathcal{H}^0$. This starts the induction.

Now suppose that N intersects $\mathcal{F}^1$. If N violates condition (vi), then it must have at least two arcs of intersection with ∂M. This gives it non-negative area. Also, if its area is zero, then these are the only arcs of intersection with ∂M, and also N is disjoint from $\mathcal{F}^1$. But we have already dealt with this case.

If N violates condition (v) of the definition of normality, it contains two sub-arcs in the same 1-handle of $\mathcal{F}$. Extend these to maximal sub-arcs A_1 and A_2 of N, with the property that each A_i lies in $\mathcal{F}$ and that A_1 and A_2 run along the same handles of $\mathcal{F}$. Since we are assuming that N has non-positive area and that it runs over some 1-handle of $\mathcal{F}$, it therefore intersects ∂M in at most one arc. Hence, at least one arc A_3 of cl$(N - (A_1 \cup A_2))$ lies in $\mathcal{F}$. If ∂A_3 lies at the same ends of A_1 and A_2, then A_3 can be closed up to form an admissible curve. If ∂A_3 lies at

opposite ends of A_1 and A_2, then the $A_1 \cup A_3$ can be closed to form an admissible curve. In either case, the inductive hypothesis gives that the combinatorial area of this admissible curve is non-negative and hence the combinatorial area of N is positive.

If N is not embedded but does not violate (v) or (vi), then let P_1 and P_2 be points on N which are coincident in $\partial \mathcal{H}^0$. Then P_1 and P_2 lie in some 0-handle of $\mathcal{F}$. If both arcs of $N - (P_1 \cup P_2)$ meet ∂M, then, since N has non-positive area, it is disjoint from $\mathcal{F}^1$. We have proved the lemma in this case. Suppose therefore that at least one arc of $N - (P_1 \cup P_2)$ lies entirely in $\mathcal{F}$. Join the endpoints of this arc to form a closed curve N_1. Then N_1 satisfies all the conditions of normality, except that it may be non-embedded. However, N_1 has fewer self-intersections than N, and so we may repeat this process and end with a normal curve. This has non-negative area and so N has positive area. $\square$

The following result is analogous to the Gauss-Bonnet theorem.

Proposition 4.3. *Let F be an admissible surface in a handle structure $\mathcal{H}$ arising from an angled spine. Let* $\mathrm{Length}(\partial F - \partial M)$ *be the number of arcs of intersection between $\partial F - \partial M$ and the 1-handles of $\mathcal{H}$. Then the combinatorial area of F satisfies* $a(F) = -2\pi\chi(F) + 2\pi\mathrm{Length}(\partial F - \partial M)$.

Proof. This is a very straightforward counting argument. Examine the formula for combinatorial area, term by term. Summing the (-2π) terms for each 0-handle of F gives $-2\pi|\mathcal{H}^0 \cap F|$. Each exterior angle is π minus the corresponding interior angle. The interior angles sum to $2\pi|\mathcal{H}^2 \cap F|$. We need to know the total number $|F \cap \mathcal{F}^1|$ of exterior angles of F. Arranged around each 0-handle H of F are components of $\partial H \cap \mathcal{F}^0$, alternating with components of $\partial H \cap \mathcal{F}^1$, $\partial H \cap \partial M$ and $\partial H - \partial \mathcal{H}^0$. Each component of $F \cap \mathcal{H}^1$ gives two components of $F \cap \mathcal{F}^0$. Hence,

$$2|\mathcal{H}^1 \cap F| = |F \cap \mathcal{F}^0| = |F \cap \mathcal{F}^1| + |F \cap \partial\mathcal{H}^0 \cap \partial M| + \mathrm{Length}(\partial F - \partial M).$$

Therefore, $\pi|F \cap \mathcal{F}^1|$, together with the sum of the $3\pi|\partial H - \partial\mathcal{H}^0|$ and $\pi|H \cap \partial M|$ terms gives $2\pi|\mathcal{H}^1 \cap F| + 2\pi\mathrm{Length}(\partial F - \partial M)$. $\square$

Casson used this to show the following result.

Proposition 4.4. *An orientable 3-manifold M with an angled ideal triangulation contains no normal 2-spheres and the only normal tori arise as links of the ideal vertices. Also, each boundary component of M is a torus.*

Proof. Let F be a normal 2-sphere or torus. Then the combinatorial area of each 0-handle of F is non-negative, and so the combinatorial area of F is non-negative. But, by Proposition 4.3, the combinatorial area of a 2-sphere is -4π and that of a torus is zero. Hence, we can have no normal 2-sphere. Each 0-handle of a normal torus has zero combinatorial area, and therefore, by the remarks after Lemma 4.1, is a triangle as in Figure 13. These triangles join to form the link of an ideal vertex. Furthermore, the link of an ideal vertex is comprised of triangles, each with zero combinatorial area. Hence, by Proposition 4.3, each boundary component of M has zero Euler characteristic, and so is a torus. □

In a similar spirit, we can restrict the possible surfaces with non-negative Euler characteristic in an angled spine.

Proposition 4.5. *Let M be a 3-manifold with a handle structure $\mathcal{H}$ arising from an angled spine. Then $\mathcal{H}$ contains no normal properly embedded spheres or discs. Also, if the angled spine is dual to an angled ideal triangulation, then each normal annulus properly embedded in M is constructed by picking a 2-handle H_2 of $\mathcal{H}$, and then taking $\mathrm{cl}(\partial\mathcal{N}(H_2)-\partial M)$, where $\mathcal{N}(H_2)$ is a small regular neighbourhood of H_2.*

Proof. As explained above, the combinatorial area of a normal surface in non-negative. However, by Proposition 4.3, the combinatorial area of a properly embedded disc or sphere is negative. Thus, $\mathcal{H}$ can contain no normal spheres or properly embedded discs. Similarly, the combinatorial area of a properly embedded annulus F is zero. So, if $\mathcal{H}$ is dual to an angled ideal triangulation, each 0-handle of F is either a triangle or a boundary bigon. At least one 0-handle of F is a boundary bigon. Also, a boundary bigon and a triangle cannot be adjacent 0-handles in F, and therefore, every 0-handle of F is a boundary bigon. These combine to form an annulus as described in the proposition. □

Corollary 4.6. *An orientable 3-manifold with an angled ideal triangulation is irreducible, atoroidal and not Seifert fibred.*

Proof. Irreducibility and atoroidality follow from Proposition 4.4. Seifert fibred manifolds with non-empty boundary are either the solid torus or have essential properly embedded annuli. In the former case, there is a meridian disc in normal form. In the latter case, we may find an essential properly embedded annulus in normal form. This contradicts Proposition 4.5. □

We cannot in general deduce information about annuli and tori in angled spines, since there may exist many types of normal curves in $\partial\mathcal{H}^0$ with zero combinatorial area. The following result asserts that manifolds with angled spines are fairly ubiquitous.

Theorem 4.7. *Let M be a compact orientable irreducible atoroidal 3-manifold with non-empty boundary. Suppose that M contains no properly embedded essential discs or annuli, and that M is not a 3-ball. Then M has an angled spine.*

Proof. Let us first deal with the case where each boundary component of M is a torus. Then, by [24], $M - \partial M$ has a complete finite volume hyperbolic structure. By [7], $M - \partial M$ admits an expression as a union of convex hyperbolic ideal polyhedra with faces identified isometrically in pairs. The dual of this is a spine, with an interior angle associated with each component of $\mathcal{F}^1$, arising from the interior dihedral angle of the corresponding edge in one of the polyhedra. We claim that this is actually an angled spine. As in the proof of Lemma 4.1, the second part of the definition of an angled spine is immediately verified. To check the first part, consider a normal curve N in $\partial\mathcal{H}^0$. We need to show that $a(N) \geq 0$. If N lies entirely in $\mathcal{F}$, then a result of Rivin (Theorem 1 of [21]) about convex hyperbolic ideal polyhedra implies that the sum of the exterior angles of N is at least 2π, and so $a(N) \geq 0$. However, N may intersect ∂M. If it intersects ∂M in at least two arcs, then $a(N) \geq 0$. Suppose that N intersects ∂M in precisely one arc A. If we replace A with an arc A' in $\mathcal{F}$ skirting around the component of $\partial\mathcal{H}^0 - \mathcal{F}$ containing A, the result is an admissible curve N' in $\mathcal{F}$. There are two choices A'_1 and A'_2 for A' corresponding to the two ways of going around the component of $\partial\mathcal{H}^0 - \mathcal{F}$. The exterior angles of A'_1 and A'_2 sum to 2π, and hence we can ensure that the exterior angles of A' sum to at most π. Hence, $a(N) \geq a(N')$.

Now, the curve N' might not be normal, but by the argument of Lemma 4.2, $a(N')$ is at least that of a normal curve in $\mathcal{F}$, and so $a(N') \geq 0$. This verifies that we have an angled spine.

Consider now the case where M contains some non-toral boundary components. Let T be the toral boundary components of M (possibly, $T = \emptyset$). Let Y be the 3-manifold obtained by doubling $M - T$ along $\partial M - T$. Then our assumptions about M imply that Y is irreducible and atoroidal. Now, Y is Haken, since $\partial M - T$ is incompressible. Hence, Y is not a Seifert fibre space, since otherwise, according to VI.34 of [14], $\partial M - T$ would separate Y into two I-bundles over compact surfaces, which would imply that M contained a properly embedded essential annulus, contrary to assumption. Hence, by [24], Y has a complete finite volume hyperbolic structure. It admits an isometric involution τ which swaps the two copies of $M - T$. The fixed point set of τ is a totally geodesic copy of $\partial M - T$. Hence, we have the well-known result that M admits a complete hyperbolic structure with $\partial M - T$ being totally geodesic.

In the case where $T \neq \emptyset$, Y admits a canonical expression as a union of convex hyperbolic ideal polyhedra with faces identified in pairs [7]. This is preserved by τ. Hence, the intersection of these polyhedra with one half of Y then gives an expression of $M - T$ as a union of convex hyperbolic polyhedra P (with some vertices possibly at infinity), with $\partial M - T$ being a union of faces of P. Each edge touching $\partial M - T$ either lies entirely in $\partial M - T$ or is at right-angles to it.

Similarly, when $T = \emptyset$, according to [16], M admits such a representation as a union of convex hyperbolic polyhedra P, with each edge touching $\partial M - T$ either lying entirely in $\partial M - T$ or at right-angles to it. These polyhedra are known as 'truncated polyhedra' since they come from hyper-ideal polyhedra by performing a perpendicular truncation of the infinite-volume ends.

The angled spine of M is obtained by dualising P in the following fashion: associate a 0-handle of the spine with each polyhedron of P, associate a 1-handle of the spine with each face of $P - \partial M$, and associate a 2-handle of the spine with each edge of P not lying entirely in ∂M. Again, to show that this is an angled spine, we must show that each normal curve N in $\partial \mathcal{H}^0$ has non-negative combinatorial area. As in the case above where P is ideal, we can assume that N

misses T (but it might run over components of $\partial M - T$). Hence, it corresponds to a normal curve (also called N) in boundary of a polyhedron P' of P which is transverse to the edges of P'. Each arc of intersection with $\partial M - T$ contributes two external angles of $\pi/2$ to N. Hence, we may assume that N has at most one arc of intersection with $\partial M - T$. If it misses $\partial M - T$, then it corresponds to a path N_1 in the boundary of the hyper-ideal polyhedron P_1 obtained by re-attaching an infinite volume end to each component of $P' \cap (\partial M - T)$. If N hits $\partial M - T$, then let P_2 be the hyper-ideal polyhedron obtained by doubling P' along the component of $P' \cap (\partial M - T)$ intersecting N, and then attaching infinite volume ends to the remaining components of $P' \cap (\partial M - T)$. Two copies of N (one in each half of P_2) join to form a closed curve N_2 in the boundary of P_2, with $a(N_2) = 2a(N)$. By the argument of [21] generalised to hyper-ideal polyhedra, N_1 and N_2 have non-negative combinatorial area. Hence, so does N. $\square$

The combinatorial 2π theorem outlined in the introduction will be proved using results about admissible surfaces with boundary. Let $\mathcal{H}$ be the handle structure arising from an angled spine. We wish to define the combinatorial length of a slope s on ∂M. Pick a curve C representing a non-zero multiple of s which respects the handle structure on ∂M, in the sense that it is disjoint from the 2-handles and is vertical in the 1-handles. Let $C_1^0, \ldots, C_n^0$ be the sequence of arcs of intersection between C and the 0-handles of ∂M, and let $C_1^1, \ldots, C_n^1$ be the sequence of arcs of intersection between C and the 1-handles of ∂M, where C_i^1 lies between C_i^0 and C_{i+1}^0. We view C_{n+i}^0 as C_i^0, and C_{n+i}^1 as C_i^1. Define an *inward extension* of C to be a surface $\mathcal{E}$ lying in M, comprised of handles $H_1^0, \ldots, H_n^0, H_1^1, \ldots, H_n^1$, such that

- each H_i^j is a disc in a j-handle of $\mathcal{H}$;

- each H_i^j contains C_i^j;

- each H_i^1 respects the product structure of the 1-handle of $\mathcal{H}$ in which it lies;

- each intersection $H_i^0 \cap \partial \mathcal{H}^0$ is admissible;

- the arcs of $\mathcal{F}^0 \cap H_i^0$ and $\mathcal{F}^0 \cap H_i^1$ touching $C_i^0 \cap C_i^1$ agree;

- the arcs of $\mathcal{F}^0 \cap H_i^1$ and $\mathcal{F}^0 \cap H_{i+1}^0$ touching $C_i^1 \cap C_{i+1}^0$ agree.

Define the *weight* of $\mathcal{E}$ to be

$$w(\mathcal{E}) = \sum_{i=1}^{n} a(H_i^0)/|H_i^0 \cap \partial M|.$$

Define the *combinatorial length* of C to be

$$l(C) = \inf\{w(\mathcal{E}) : \mathcal{E} \text{ is an inward extension of } C\},$$

and define $l(C)$ as infinite if C has no inward extension. This happens if an arc C_i^0 has endpoints in the same component of $\partial\mathcal{F}^0 - \mathcal{F}^1$, for then a curve ∂H_i^0 containing C_i^0 must fail condition (iv) in the definition of normality, and so cannot be admissible. In fact, if no arc C_i^0 is of this form, then C has an inward extension which is an annulus formed by pushing C a little into the interior of M. Let the *combinatorial length* of s be

$$l(s) = \inf\{l(C) : C \text{ is a curve representing a non-zero multiple of } s\}.$$

These definitions have been designed specifically so that the following proposition holds.

Proposition 4.8. *Let F be an admissible surface in a handle structure on M arising from an angled spine. Let $C_1, \ldots, C_m$ be the components of $\partial F \cap \partial M$, each C_j representing a non-zero multiple of some slope $s_{i(j)}$. Then*

$$a(F) \geq \sum_{j=1}^{m} l(s_{i(j)}).$$

Proof. The handles of F touching C_j form an inward extension of C_j. Summing the weights of these inward extensions over all C_j gives the required inequality. $\square$

The following result is our 'word hyperbolic Dehn surgery theorem' for angled spines. Note that this, together with Corollary 4.6, gives the version of Theorem 4.9 mentioned in the introduction.

Theorem 4.9. *Let M be a compact orientable 3-manifold with an angled spine. Suppose that M is atoroidal and not a Seifert fibre space, and has boundary a non-empty union of tori. Let $s_1, \ldots, s_n$ be a collection of slopes on ∂M, with one*

s_i on each component of ∂M, and each with combinatorial length more than 2π. Then the manifold obtained by Dehn filling M along these slopes is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.

Proof. Note that M is irreducible, since it has an angled spine. Since we are assuming that it is atoroidal and not a Seifert fibre space, its interior supports a complete finite volume hyperbolic structure [24]. Suppose first that $M(s_1, \ldots, s_n)$ is reducible or toroidal. Then we can find F a punctured sphere or torus in M, with F incompressible and ∂-incompressible in M. Hence, F can be ambient isotoped into normal form in $\mathcal{H}$, the thickened angled spine of M. Let $C_1, \ldots, C_{|\partial F|}$ be the boundary components of F. We can ensure that C_j is essential in ∂M having slope $s_{i(j)} \in \{s_1, \ldots, s_n\}$. Proposition 4.3 gives that

$$a(F) = -2\pi\chi(F) \leq 2\pi|\partial F|,$$

but Proposition 4.8 gives that

$$a(F) \geq \sum_{j=1}^{|\partial F|} l(s_{i(j)}) > 2\pi|\partial F|,$$

which together give a contradiction. A similar argument gives that the core of each surgery solid torus in $M(s_1, \ldots, s_n)$ has infinite order in $\pi_1(M(s_1, \ldots, s_n))$.

To show the word hyperbolicity of $\pi_1(M(s_1, \ldots, s_n))$, we will use Gabai's Ubiquity theorem. Let K be a curve in M which is homotopically trivial in $M(s_1, \ldots, s_n)$. Assume that K is disjoint from the 2-handles of $\mathcal{H}$ and respects the product structure on the 1-handles. Then, there is a compact planar surface F in M, with ∂F being K together with curves $C_1, \ldots, C_{|F \cap \partial M|}$ on ∂M, each C_j being a non-zero multiple k_j of some slope $s_{i(j)}$. By Lemma 2.5, we may assume that F is homotopically incompressible and homotopically ∂-incompressible. We may therefore homotope F (keeping K invariant) to an admissible surface. Let Length(K) be the number of arcs of intersection between K and $\mathcal{H}^1$, and let $\epsilon > 0$ be a real number with $l(s_i) \geq 2\pi + \epsilon$ for all i. Then Propositions 4.8 and 4.3 give

that

$$(2\pi + \epsilon)|F \cap \partial M| \leq \sum_{j=1}^{|F \cap \partial M|} l(s_{i(j)})$$

$$\leq a(F) = -2\pi\chi(F) + 2\pi\text{Length}(K)$$

$$< 2\pi|F \cap \partial M| + 2\pi\text{Length}(K),$$

from which we get

$$|F \cap \partial M| < (2\pi/\epsilon)\text{Length}(K).$$

Theorem 2.1 gives us that $\pi_1(M(s_1, \ldots, s_n))$ is word hyperbolic. $\square$

The conclusions of Theorem 4.9 still hold (with the exception of word hyperbolicity) if one leaves some boundary components of M unfilled. Also, one can allow M to have some non-toral boundary components.

Of course, the above theorem begs the question of which slopes on ∂M have combinatorial length more than 2π. The purpose of the following proposition is to give a simple lower bound in the case of angled ideal triangulations.

Proposition 4.10. *Let M be an orientable 3-manifold with an angled ideal triangulation. Define the length of each edge E of the associated triangulation of ∂M to be $\min\{\alpha_1, \ldots, \alpha_6\}/2$, where $\alpha_1, \ldots, \alpha_6$ are the interior angles of the two triangles adjacent to E. Each simplicial curve in the 1-skeleton of ∂M then inherits a length. The shortest length of a curve of slope s is a lower bound for the combinatorial length $l(s)$.*

Proof. The ideal triangulation of M dualises to a special spine, and so the boundary of M inherits a handle structure in which each 0-handle has valence three. Consider a curve C representing a non-zero multiple of the slope s in ∂M, such that C is disjoint from the 2-handles and respects the product structures on the 1-handles. As above, let $C_1^1, \ldots, C_n^1$ be the arcs of intersection between C and the 1-handles of ∂M. For each C_i^1, let E_i be the associated edge in the dual triangulation. Let $\mathcal{E}$ be an inward extension of C.

We term C_i^1 a *hugging arc* if C_{i-1}^1, C_i^1 and C_{i+1}^1 are all adjacent to some 2-handle of ∂M and are arranged in order around that 2-handle, as shown in Figure 15. The point behind this definition is that if H_i^0 and H_{i+1}^0 are both boundary bigons in the inward extension $\mathcal{E}$, then C_i^1 must be a hugging arc.
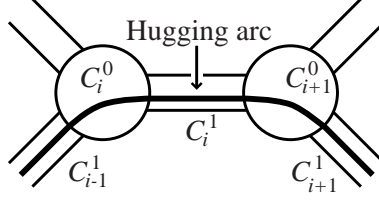
Figure 15.

Claim. Suppose that $H = H_i^0$ is not a boundary bigon. Let $\{\alpha_1, \alpha_2, \alpha_3\}$ be the interior angles of the 0-handle of $\mathcal{H}$ containing H. Then $a(H)/|H \cap \partial M| \geq \min\{\alpha_1, \alpha_2, \alpha_3\}$.

Let $\epsilon_1, \ldots, \epsilon_t$ be the exterior angles inherited by ∂H. If ∂H intersects ∂M in at least three arcs, then

$$\frac{a(H)}{|H \cap \partial M|} = \frac{(\sum_{j=1}^t \epsilon_j) - 2\pi + \pi|H \cap \partial M| + 3\pi|\partial H - \partial \mathcal{H}^0|}{|H \cap \partial M|}$$

$$\geq \pi\left(1 - \frac{2}{|H \cap \partial M|}\right) \geq \pi/3.$$

But $\alpha_1 + \alpha_2 + \alpha_3 = \pi$, and therefore $\min\{\alpha_1, \alpha_2, \alpha_3\} \leq \pi/3$. Therefore, in the case where $|H \cap \partial M| \geq 3$, the claim is proved. For $|H \cap \partial M| = 2$, we have

$$\frac{a(H)}{|H \cap \partial M|} = \frac{(\sum_{j=1}^t \epsilon_i) - 2\pi + \pi|H \cap \partial M| + 3\pi|\partial H - \partial \mathcal{H}^0|}{|H \cap \partial M|}$$

$$\geq \min\{\pi - \alpha_1, \pi - \alpha_2, \pi - \alpha_3, 3\pi\}/2$$

$$\geq \min\{\alpha_1, \alpha_2, \alpha_3\}.$$

For $|H \cap \partial M| = 1$ and ∂H not normal, then the proof of Lemma 4.2 gives that the combinatorial area of H is at least π. The normal curve N with $|N \cap \partial M| = 1$ is given in Figure 13, and its combinatorial area is at least $\min\{\alpha_1, \alpha_2, \alpha_3\}$. This proves the claim.

We will show that C naturally determines a curve C' in the 1-skeleton of the dual triangulation, in the following manner. Keeping the midpoints of C_i^1 and C_{i+1}^1 fixed, homotope the arc of C lying between them, so that it runs along E_i and then back up E_{i+1}. (We can do this because each 0-handle of the handle structure on ∂M has valence three.) The curve which we have constructed lies in the 1-skeleton of the triangulation of ∂M, but need not be simplicial, since the vertex joining E_{i-1} to E_i may be the same as the vertex joining E_i to E_{i+1}. This

36

happens precisely when C_i^1 is a hugging arc. However, a further homotopy creates a simplicial curve C', with each edge E_i of C' being dual to a non-hugging arc C_i^1. Let $\{\alpha_1, \ldots, \alpha_6\}$ be the six interior angles of the two triangles adjacent to E_i. Then, at least one of ∂H_i^0 and ∂H_{i+1}^0 is not a boundary bigon. Therefore, by the claim,

$$\left(\frac{a(H_i^0)}{|H_i^0 \cap \partial M|} + \frac{a(H_{i+1}^0)}{|H_{i+1}^0 \cap \partial M|} \right) /2 \geq \min\{\alpha_1, \ldots, \alpha_6\}/2.$$

Summing this from $i = 1$ to n gives that the combinatorial length of C is at least the length of C' in the path metric on the 1-skeleton of the triangulation of ∂M. Since C' represents a non-zero multiple of the slope s, it is possible to perform a cut-and-paste on C' in the vertices of ∂M, making it into a collection of curves, each with slope s. In particular, the length of C' is at least the length of the shortest curve in the 1-skeleton of ∂M with slope s. $\square$

Note that the factor $1/2$ arises in the definition of edge length in the triangulation of ∂M to avoid double counting. This is because we may have successive edges E_i and E_{i+1} in C', dual to non-hugging arcs C_i^1 and C_{i+1}^1, but with ∂H_i^0 and ∂H_{i+2}^0 both boundary bigons. Therefore, we must share the combinatorial area of H_{i+1}^0 out between E_i and E_{i+1}. However, if successive edges E and E' of C' are not dual to successive edges of C (because there is some hugging arc between them), then this double counting cannot occur. Hence, the lower bound in Proposition 4.10 need not be sharp. Nevertheless, Proposition 4.10 has the following immediate corollary.

Corollary 4.11. *Let M be an orientable 3-manifold with an angled ideal triangulation. Then at most finitely many slopes on each component of ∂M have combinatorial length no more than 2π.*

5. Surgery along alternating links

Our goal here is to use the techniques of the previous section to deduce the following word hyperbolic Dehn surgery theorem for alternating links.

Theorem 5.1. *Let D be a connected prime alternating diagram of a link L in S^3. For each component K of L, pick a surgery coefficient p/q (in its lowest terms) with $|q| > 8/t(K, D)$. The manifold obtained by Dehn surgery along L, via these surgery coefficients, is irreducible, atoroidal and not a Seifert fibre space, and has infinite, word hyperbolic fundamental group.*

The number $t(K, D)$ was defined in the introduction to be half the number of non-bigon edges of $G(D)$ lying in K. Similarly, the twist number $t(D)$ was defined to be half the number of non-bigon edges in $G(D)$. It has the following simple interpretation. Define a *twist* of D to be a maximal connected subgraph of $G(D)$ with edges comprised of bigon edges. A single vertex of $G(D)$ incident to no bigon edge is viewed as a twist. Twists in a connected diagram D can be of several forms: either that shown in Figure 16; or they can comprise the whole diagram, in which case D is the standard diagram of the $(2, n)$-torus link for some n; or they can form a Hopf link summand with two crossings. Henceforth, suppose that D has positive twist number and has no Hopf link summands with two crossings. Then, it is not hard to see that $t(D)$ is simply the number of twists in D.

A twist

Figure 16.

If one replaces each twist of $G(D)$ with a single vertex, then one obtains a new 4-valent graph G_1 with precisely $t(D)$ vertices. One can then reconstruct $G(D)$ from G_1 by replacing some or all of its vertices with twists. This procedure can be effected in a surgical manner, as follows. Start with G_1 and replace some or all of its vertices with two vertices joined by a bigon region, creating a new 4-valent graph G_2. Then assign under-over crossing information to G_2 creating an alternating link L_2, say. Add unknotted components to L_2, one component for

each vertex of G_1, which encircles the crossing (or pair of crossings) as shown in Figure 17. Let L_3 be the resulting link. It is an augmented alternating link, in the sense of [1]. Then the link L is obtained from L_3, by performing $1/q$ surgery on each of the new unknotted components (where $q \in \mathbb{Z}$ depends on the component of L_3).
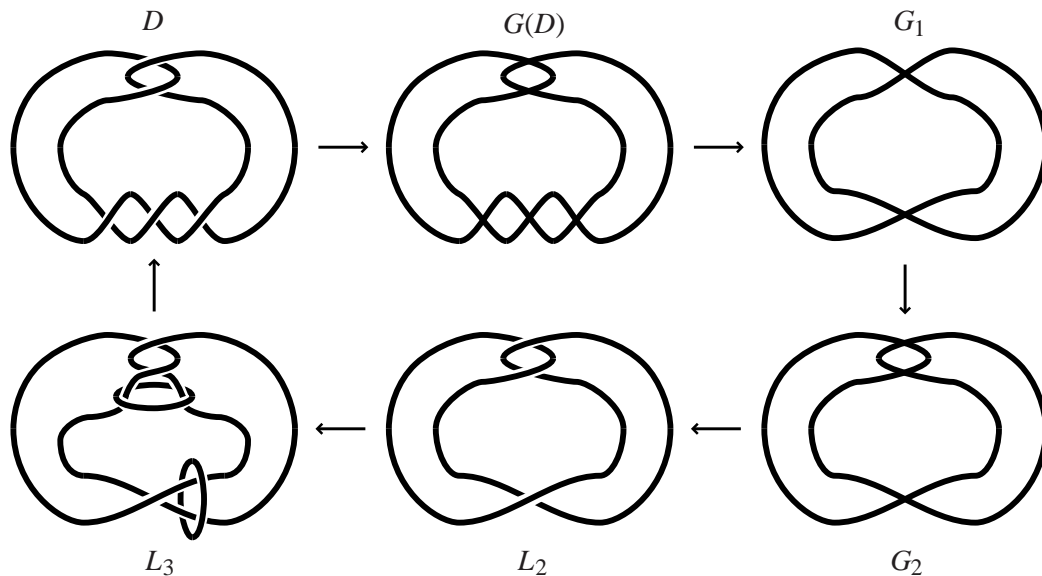

Figure 17.

For a fixed integer n, there is only a finite number of 4-valent planar graphs G_1 with n vertices. At each stage of the construction of L_3 from G_1, we had only a finite number of choices. Thus, we have shown that the alternating links with a given twist number n are all obtained from a finite list of augmented alternating links via surgery on some of the components of the link. Adams [1] showed that the complement of an augmented alternating link L_3 is hyperbolic. By [23], if the complement of the alternating link L is hyperbolic, then its volume is at most that of the complement of L_3. Applying this argument to knots with twist number at most 8, we have the following corollary to Theorem 5.1.

Corollary 5.2. *There is a real number c with the property that if K is an alternating knot whose complement has a complete hyperbolic structure with volume at least c, then every non-trivial surgery on K yields a 3-manifold which is irreducible, atoroidal and not Seifert fibred, and has infinite, word hyperbolic fundamental group.*

It has been known for several years that the complement of a non-split alternating link can be expressed as the union of two ideal polyhedra glued along their faces. Dually, a non-split alternating link complement has a spine with some nice properties. We briefly recall here the salient points of the theory. A more complete description can be found in [3] and [17].

The alternating link diagram D is viewed as lying on a 2-sphere S^2 embedded in S^3. The link itself is viewed as lying in S^2, except near each crossing, where two arcs skirt above and below S^2. The spine of the alternating link exterior has two 0-cells, one lying in each of the two 3-balls $S^3 - \text{int}(\mathcal{N}(S^2))$; it has a 1-cell for each region of the diagram, running between the two 0-cells; and it has a 4-valent 2-cell for each crossing. Let $\mathcal{H}$ be the handle structure arising by thickening the spine. The surface $\mathcal{F} = \mathcal{H}^0 \cap (\mathcal{H}^1 \cup \mathcal{H}^2)$ has two components, one in each 0-handle of $\mathcal{H}$. It is a crucial observation that each component of $\mathcal{F}$ is dual to the underlying graph $G(D)$ of the diagram. Thus, dually, the link complement is obtained by gluing together two ideal polyhedra P_1 and P_2, each with boundary graph a copy of $G(D)$.
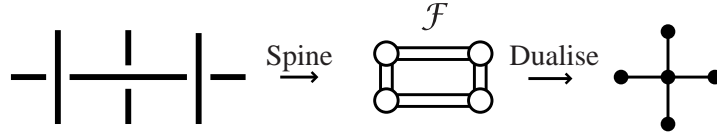

Figure 18.

Each face of ∂P_1 is glued to the corresponding face of ∂P_2, but with a twist. The orientation of the twist is determined as follows. The complimentary regions of $G(D)$ admit a black-and-white 'checkerboard' colouring, with adjacent regions being coloured with different colours. Then, we attach the black (respectively, white) regions of ∂P_1 to the corresponding region of ∂P_2 with a single clockwise (respectively, anti-clockwise) twist. This was first observed by Thurston [23] in the case of the Borromean rings, where he likened this gluing procedure to the gears of a machine. An example is given in Figure 19.
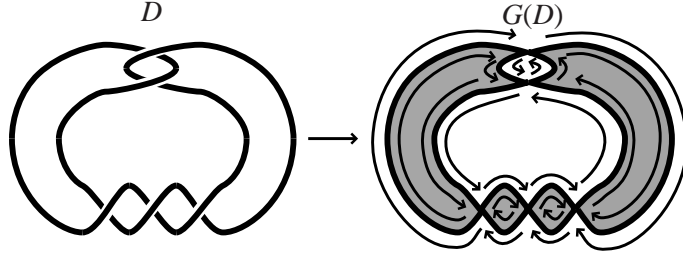
Figure 19.

When the exterior M of a non-split alternating link L is given this handle structure, its boundary inherits a particularly simple handle structure. The simplest way to see this is to embed ∂M as a normal surface. We then see that each component T of ∂M has the form of Figure 20. Each 0-handle of T has valence four. These 0-handles are arranged along T, alternating between 0-handles above S^2 and 0-handles below S^2. If C is a meridian curve in ∂M encircling an edge E of $G(D)$, then C intersects two 1-handles and two 0-handles of ∂M. We say that these handles are *associated* with the edge E. The two 1-handles associated with E each join the 0-handles associated with E. Hence, if K is some component of L, the total number of 0-handles in $\partial \mathcal{N}(K)$ is equal to the number of edges $e(K, D)$ of K in $G(D)$.
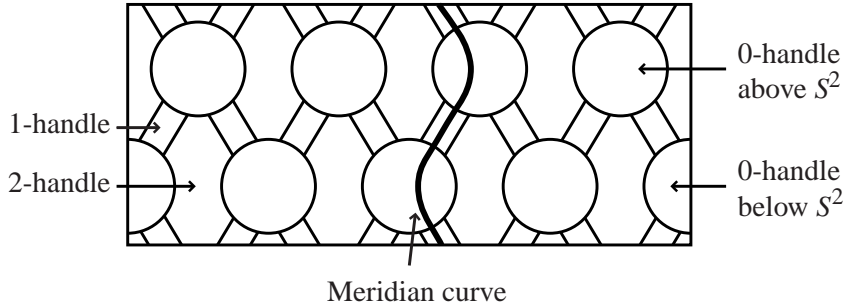

Figure 20.

In order to utilise the results of Section 4, we need to place an angled structure on the spine. The simplest way to do this is to assign an 'exterior angle' in the range $(0, \pi)$ to each edge of the graph $G(D)$. Since each 1-handle of $\mathcal{F}$ is dual to some edge of $G(D)$, this assigns an exterior angle to that 1-handle. Of course, one must check that these angles do genuinely determine an angled spine.

Recall that a diagram D is *reduced* if every crossing of D is adjacent to four

distinct complimentary regions of $G(D)$. Note that if a diagram is prime and has more than one crossing, then it is reduced.

Proposition 5.3. *Assigning an exterior angle to each edge of $G(D)$ gives an angled spine if and only if each of the following conditions are satisfied:*

(i) *D is reduced;*

(ii) *the exterior angles of the edges around any vertex of $G(D)$ sum to 2π;*

(iii) *any simple closed curve in S^2 which intersects $G(D)$ transversely in the interior of edges of $G(D)$ and touches each edge at most once, has total exterior angle at least 2π.*

Proof. The verification of the first part of the definition of an angled spine is a simple argument analogous to the first half of the proof of Theorem 4.7, and is omitted. To verify the second part, recall that a 2-handle of $\mathcal{H}$ is associated with each crossing of D. The four exterior angles assigned to the edges of $G(D)$ emanating from that crossing are precisely the four exterior angles assigned to the 2-handle. Hence, by (ii), the four exterior angles around that 2-handle sum to 2π, and therefore so do the four interior angles. $\square$

In Figure 21, we assign exterior angles to the standard diagrams of the figure-eight knot, the Borromean rings and the Whitehead link. The angles satisfy the conditions of Proposition 5.3 and so define angled spines on the link exteriors. In this case, our choice of angles has been inspired by the hyperbolic structures on these link complements.
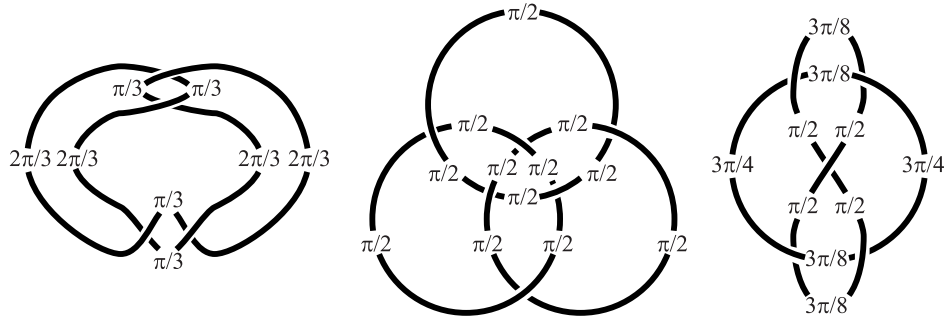

Figure 21.

It is the example of the Borromean rings which readily generalises to *any*

connected prime alternating diagram D with more than one crossing. If we assign an angle $\pi/2$ to each edge of $G(D)$, then the conditions of Proposition 5.3 are satisfied. We will call this the *canonical angled spine* arising from D.

The major technical result of this section is the following.

Theorem 5.4. *Let L be a link with a connected prime alternating diagram D having more than one crossing. Give the exterior of L the canonical angled spine arising from D. Then the combinatorial length of the slope p/q on a component K of L is at least $|q|\,t(K,D)\,\pi/4$.*

Proof of Theorem 5.1 from Theorem 5.4. If D has precisely one crossing, then its twist number is zero, and the theorem is trivial. So suppose that D has more than one crossing. Menasco [18] proved that the exterior of L is irreducible, atoroidal and not a Seifert fibre space, unless D is the standard diagram of a $(2,n)$-torus link for some n. In this case, $t(K,D) = 0$, and the statement of the theorem is empty. Otherwise, Theorem 5.1 follows from Theorem 5.4 and Theorem 4.9. □

Let M be the exterior of L. Let C be a curve on $\partial\mathcal{N}(K)$ representing a non-zero multiple k of the slope p/q. As before, we let $C_1^0, \ldots, C_n^0$ be the sequence of arcs of intersection between C and the 0-handles of ∂M, and let $C_1^1, \ldots, C_n^1$ be the sequence of arcs of intersection between C and the 1-handles of ∂M, where C_i^1 lies between C_i^0 and C_{i+1}^0. Let $\mathcal{E} = H_1^0 \cup \ldots \cup H_n^0 \cup H_1^1 \cup \ldots \cup H_n^1$ be an inward extension of C.

Let $Y \to \partial\mathcal{N}(K)$ be the cover corresponding to the subgroup of $\pi_1(\partial\mathcal{N}(K))$ generated by the meridian. In other words, unwind $\partial\mathcal{N}(K)$ in the longitudinal direction. We assign an x co-ordinate to each 0-handle of Y, with adjacent 0-handles having co-ordinates differing by 1. Hence, C lifts to an arc $\tilde{C}$ in Y which starts at $x = 0$ and ends at $x = k|q|e(K,D)$. Let $\tilde{C}_i^j$ be the lift of C_i^j in $\tilde{C}$. Let x_i be the x co-ordinate of $\tilde{C}_i^0$, and let $x_i' = (x_{i-1} + x_{i+1})/2$. Note that, from Figure 22, each x_i' is an integer.
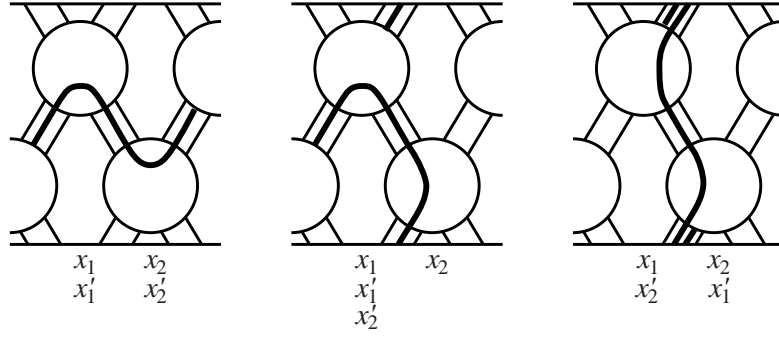
Figure 22.

Lemma 5.5. *For each integer j with $x_1' \leq j < x_1' + k|q|e(K,D)$, the 0-handles of Y at $x = j$ and $x = j+1$ have at least one arc $\tilde{C}_i^1$ running between them for which $x_i' \neq x_{i+1}'$.*

Proof. Note that $|x_{i+1}' - x_i'| \leq 1$. Also, if $x_i' \neq x_{i+1}'$, then $\{x_i', x_{i+1}'\} = \{x_i, x_{i+1}\}$. Therefore, for each integer j with $x_1' \leq j < x_1' + k|q|e(K,D)$, the pair $\{j, j+1\}$ occurs as $\{x_i', x_{i+1}'\}$ for some arc $\tilde{C}_i^1$. □

The difficulty behind Theorem 5.1 is that there exist several different normal curves N in $\partial \mathcal{H}^0$ which have zero combinatorial area. These are listed below (where we view each curve N as lying in S^2 and intersecting $G(D)$, since $\mathcal{F}$ is dual to $G(D)$).
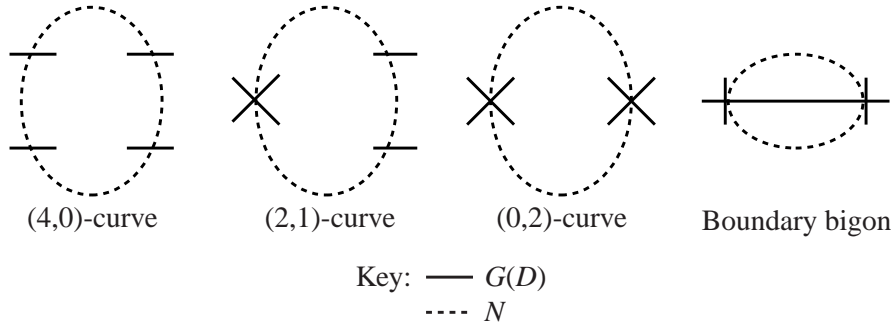

Figure 23.

We denote the normal curves N in Figure 23 as (k_1, k_2)-curves, where $k_1 = |N \cap \mathcal{F}^1|$, and $k_2 = |N \cap \partial M|$. If $k_1 = 0$ and $k_2 = 2$, there are two possible types of curve N, one of which is a boundary bigon. We refer to the other type of curve as a $(0,2)$-curve. We say that C_i^0 is a *skirting arc* if it runs between components of $\mathcal{F}^0 \cap \partial M$ which are adjacent round $\partial \mathcal{H}^0 \cap \partial M$ (see Figure 24).
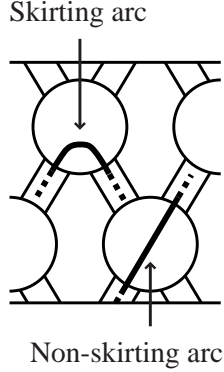
44

Skirting arc

Non-skirting arc

Figure 24.

Lemma 5.6. *If C_i^0 is a skirting arc, then H_i^0 either has positive combinatorial area or is a boundary bigon.*

Proof. Suppose that, on the contrary, H_i^0 has zero combinatorial area, but is not a boundary bigon. By Lemma 4.2, ∂H_i^0 is normal, and hence is a $(2,1)$-curve or a $(0,2)$-curve. This implies that each time ∂H_i^0 hits a vertex of $G(D)$, it crosses directly over that vertex. This contradicts our assumption that C_i^0 is a skirting arc. $\square$

Lemma 5.7. *Suppose that C_i^0 and C_{i+1}^0 are both non-skirting arcs, and that C_i^1 runs along a 1-handle of ∂M corresponding to a non-bigon edge of $G(D)$. Then at least one of ∂H_i^0 and ∂H_{i+1}^0 has positive combinatorial area.*

Proof. Suppose that, on the contrary, both ∂H_i^0 and ∂H_{i+1}^0 have zero combinatorial area. Then each is either a $(0,2)$-curve or a $(2,1)$-curve. View ∂H_i^0 and ∂H_{i+1}^0 as curves in S^2 which intersect $G(D)$. Let A_i (respectively, A_{i+1}) be the sub-arc of ∂H_i^0 (respectively, ∂H_{i+1}^0) corresponding to the component of $\partial H_i^0 \cap \partial H_i^1$ touching $C_i^0 \cap C_i^1$ (respectively, the component of $\partial H_{i+1}^0 \cap \partial H_i^1$ touching $C_{i+1}^0 \cap C_i^1$). Then A_i and A_{i+1} both lie in the same region R of $G(D)$, and they are matched onto each other by a twist which is either clockwise or anti-clockwise, depending on whether R is coloured black or white. There are a number of possibilities:

 (i) A_i runs between successive vertices of R;

 (ii) A_i runs between vertices of R which are not successive;

(iii) A_i runs between a vertex and an edge of R.

These cases are given in Figure 25. In each case, it is clear that, since D is prime and the regions of $G(D)$ admit a checkerboard colouring, the edge of $G(D)$ corresponding to C_i^1 is a bigon edge. □

Figure 25.
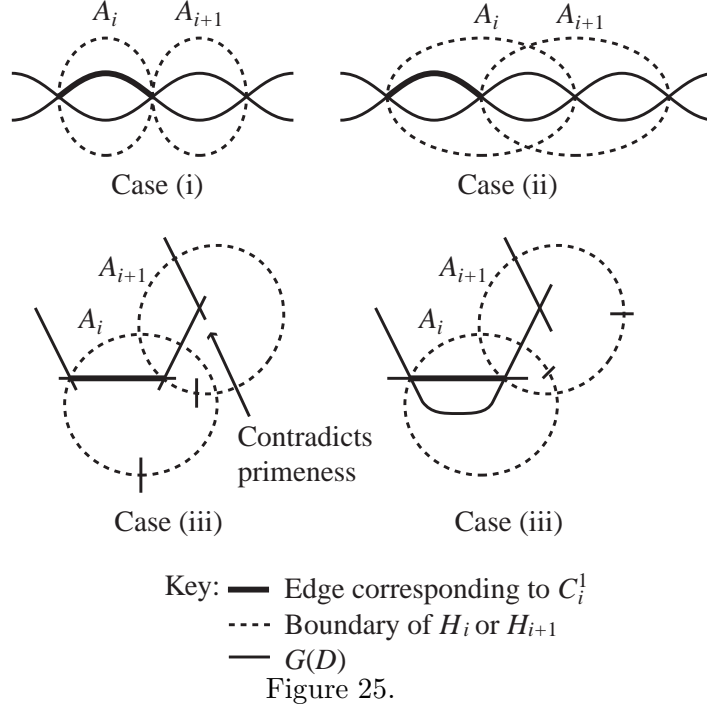
Lemma 5.8. *Suppose that precisely one of C_i^0 and C_{i+1}^0 is a skirting arc, and that C_i^1 runs along a 1-handle of ∂M corresponding to a non-bigon edge of $G(D)$. Suppose also that $x_i' \neq x_{i+1}'$. Then at least one of H_i^0 and H_{i+1}^0 has positive combinatorial area.*

Proof. Let A_i and A_{i+1} be as in the proof of Lemma 5.7, and let R be the region of $G(D)$ containing them. By Lemma 5.6, if a skirting arc does not have positive combinatorial area, then it is part of a boundary bigon. Suppose therefore that one of H_i^0 and H_{i+1}^0 is a boundary bigon. This implies that A_i and A_{i+1} both run between successive vertices of R. Suppose (for definiteness) that A_i is parallel to the edge E of $G(D)$ corresponding to C_i^1. If ∂H_i^0 is a $(0,2)$-curve, then as in case (i) of the proof of Lemma 5.7, E is a bigon edge, contrary to assumption. Hence, ∂H_i^0 must be the boundary bigon, and ∂H_{i+1}^0 must be the $(0,2)$-curve. But, then, in fact $x_i' = x_{i+1}'$, contrary to assumption. □

A simple application of the formula for combinatorial area gives the following

result. The calculations are rather similar to those in Proposition 4.10.

Lemma 5.9. *If H_i^0 has positive combinatorial area, then $a(H_i^0)/|H_i^0 \cap \partial M| \geq \pi/4$.*

The only type of normal curve N for which $a(N)/|N \cap \partial M| = \pi/4$ is a $(1,2)$-curve given in Figure 26.
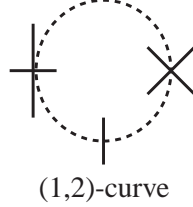

(1,2)-curve

Figure 26.

Proof of Theorem 5.1. Let C be a curve representing a non-zero multiple k of the slope p/q on $\partial \mathcal{N}(K)$, such that C is disjoint from the 2-handles and respects the product structure on the 1-handles. Define $C_1^0, \ldots, C_n^0, C_1^1, \ldots, C_n^1$ as above and let $\mathcal{E} = H_1^0 \cup \ldots \cup H_n^0 \cup H_1^1 \cup \ldots \cup H_n^1$ be an inward extension of C. Let H and H' be the 1-handles of $\partial \mathcal{N}(K)$ associated with a non-bigon edge of K. There are $2t(K,D)$ such non-bigon edges. By Lemma 5.5, there are at least $k|q|$ arcs C_i^1 on $H \cup H'$ with $x_i' \neq x_{i+1}'$. Let C_i^1 be one such arc. Then H_i^0 and H_{i+1}^0 cannot both be boundary bigons. Lemmas 5.6, 5.7 and 5.8 imply that at least one of H_i^0 and H_{i+1}^0 (H_i^0, say) has positive combinatorial area. Lemma 5.9 gives that $a(H_i^0)/|H_i^0 \cap \partial M| \geq \pi/4$. Hence,

$$w(\mathcal{E}) \geq (1/2)2t(K,D)k|q|\pi/4 \geq t(K,D)|q|\pi/4.$$

As in the discussion after Proposition 4.10, the factor $(1/2)$ is necessary to avoid double counting. Since C was an arbitrary curve representing a non-zero multiple of the slope p/q on $\partial \mathcal{N}(K)$ and $\mathcal{E}$ was an arbitrary inward extension, we have that $l(p/q) \geq |q|t(K,D)\pi/4$. $\square$

47

References

1. C. ADAMS, *Augmented alternating link complements are hyperbolic*, Low-dimensional Topology and Kleinian groups, London Math. Soc. Lecture Note Series. **112**, Cambridge Univ. Press (1986)

2. I. AGOL, *Ph.D. Thesis*, U.C. San Diego (1998)

3. I. AITCHISON, E. LUMSDEN and H. RUBINSTEIN, *Cusp structures of alternating links*, Invent. Math. **109** (1992) 473-494

4. S. BLEILER and C. HODGSON, *Spherical space forms and Dehn filling*, Topology **35** (1996) 809-833.

5. C. CAO and R. MEYERHOFF, *The orientable cusped hyperbolic manifolds of minimal volume*, Preprint.

6. C. DELMAN, *Essential laminations and surgery on 2-bridge knots*, Topology Appl. **63** (1995) 201-221.

7. D. B. A. EPSTEIN and R. C. PENNER, *Euclidean decomposition of non-compact hyperbolic manifolds*, J. Diff. Geom. **27** (1988) 67-80.

8. D. GABAI, *The ubiquitous nature of quasi-minimal semi-Euclidean laminations in 3-manifolds*, Preprint.

9. D. GABAI and W. KAZEZ, *Group negative curvature for 3-manifolds with genuine laminations*, Geom. Topol. **2** (1998) 65-77.

10. D. GABAI and U. OERTEL, *Essential laminations in 3-manifolds*, Ann. Math. **130** (1989) 41-73.

11. M. GROMOV, *Hyperbolic Groups*, Essays in Group Theory, MSRI Publ. **8**, Springer (1987) 75-263.

12. J. HEMPEL, *3-Manifolds*, Ann. of Math. Studies, No. 86, Princeton Univ. Press, Princeton, N. J. (1976)

13. C. HODGSON and S. KERCKHOFF, To appear.

14. W. JACO, *Lectures on Three-Manifold Topology*, Regional Conference Series in Mathematics, No. 43, Providence 1980, A. M. S.

15. W. JACO and U. OERTEL, *An algorithm to decide if a 3-manifold is Haken*, Topology **23** (1984) 195-209.

16. S. KOJIMA, *Polyhedral decomposition of hyperbolic 3-manifolds with totally geodesic boundary*, Aspects of Low-dimensional Manifolds, Adv. Stud. Pure Math. **20**, Kinokuniya, Tokyo (1992) 93-112.

17. W. MENASCO, *Polyhedra representation of link complements*, Low-dimensional Topology, Contemp. Math **20**, Amer. Math. Soc. (1983)

18. W. MENASCO, *Closed incompressible surfaces in alternating knot and link complements*, Topology **23** (1984) 37-44.

19. W. MENASCO and M. THISTLETHWAITE, *Surfaces with boundary in alternating knot exteriors*, J. Reine Angew. Math. **426** (1992) 47-65.

20. L. MOSHER, *Geometry of cubulated manifolds*, Topology **34** (1995) 789-814.

21. I. RIVIN, *On geometry of convex ideal polyhedra in hyperbolic 3-space*, Topology **32** (1993) 87-92.

22. R. ROBERTS, *Constructing taut foliations*, Comment. Math. Helv. **70** (1995) 516-545.

23. W. THURSTON, *The geometry and topology of three-manifolds*, Princeton (1980)

24. W. THURSTON, *Three-dimensional manifolds, Kleinian groups and hyperbolic geometry*, Bull. Amer. Math. Soc. **6** (1982) 357-381.

DPMMS

University of Cambridge

16 Mill Lane

Cambridge CB2 1SB

England.